gometría
esa el cetlation
Matemáticaicas suvererierieros:
geometria del espacio.

1 LA GEOMETRÍA DEL ESPACIO EUCLIDIANO

> Los cuaterniones vienen de Hamilton . . . y han sido maldición pura para quien, de alguna forma, los ha tocado. El vector es un sobreviviente inútil . . . y jamás ha sido de la más mínima utilidad para ningún ser viviente.
>
> Lord Kelvin

En este capítulo consideramos las operaciones básicas de los vectores en el espacio tridimensional: la suma vectorial, la multiplicación por un escalar y los productos punto y cruz. En la sección 1.5 generalizamos algunos de estos conceptos al n-espacio y revisamos las propiedades de las matrices que necesitaremos en los capítulos 2 y 3.

1.1 VECTORES EN EL ESPACIO TRIDIMENSIONAL

Los puntos P en el plano se representan mediante pares ordenados de números reales (a, b); los números a y b se llaman *coordenadas cartesianas de* P. Tracemos dos rectas perpendiculares, llamémosles ejes x y y, y bajemos perpendiculares de P a los ejes, como en la figura 1.1.1. Después de designar la intersección de los ejes x y y como origen, y de escoger unidades en estos ejes, producimos dos distancias dirigidas a y b, como se muestra en la figura; a se llama la *componente* x de P, y b se llama la *componente* y.

Los puntos en el espacio se pueden representar de manera análoga mediante ternas ordenadas de números reales. Para construir dicha representación escogemos tres rectas perpendiculares entre sí que se crucen en un punto en el espacio.

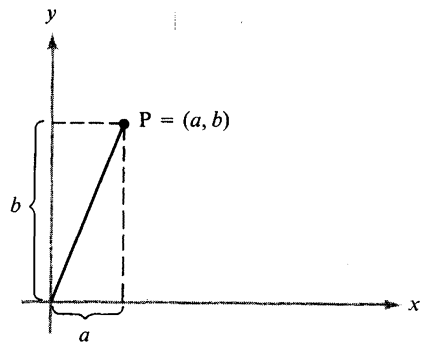


Figura 1.1.1 Coordenadas cartesianas en el plano.

Estas rectas se llaman: eje x, eje y y eje z, y el punto en el que se cruzan se llama *origen* (es nuestro punto de referencia). Escogemos una escala sobre estos ejes. Es común referirse al conjunto de ejes como *sistema de coordenadas*, y se trazan como se muestra en la figura 1.1.2.

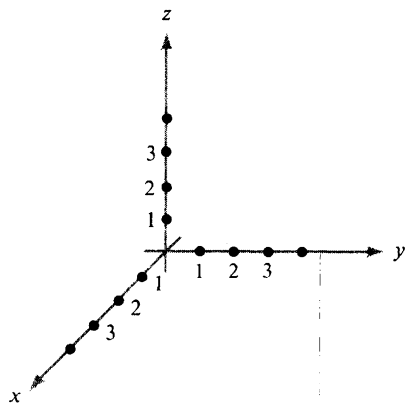


Figura 1.1.2 Coordenadas cartesianas en el espacio.

Podemos asignar a cada punto P en el espacio una terna (ordenada) única de números reales (a, b, c); y, recíprocamente, a cada terna podemos asignar un punto único en el espacio, tal y como lo hicimos para los puntos en el plano. Al origen del sistema de coordenadas le corresponde la terna $(0, 0, 0)$, y las flechas en los ejes indican las direcciones positivas. Así, por ejemplo, la terna $(2, 4, 4)$ representa un punto a 2 unidades del origen en dirección positiva a lo largo del eje x, a 4 unidades en dirección positiva a lo largo del eje y, y a 4 unidades en dirección positiva a lo largo del eje z (figura 1.1.3).

Debido a la posibilidad de asociar de esta manera los puntos del espacio con las ternas ordenadas, es común usar la expresión "punto (a, b, c)" en lugar de la

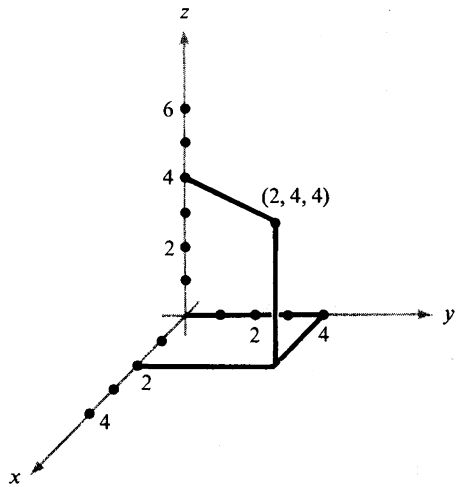


Figura 1.1.3 Representación geométrica del punto $(2, 4, 4)$ en coordenadas cartesianas.

frase más larga "punto P que corresponde a la terna (a, b, c)." Si la terna (a, b, c) corresponde a P, decimos que a es la coordenada x (o la primera coordenada), b es la coordenada y (o segunda coordenada), y c es la coordenada z (o tercera coordenada) de P. Teniendo en mente este método para representar puntos, vemos que el eje x está formado por los puntos de la forma $(a, 0, 0)$, donde a es cualquier número real; el eje y está formado por los puntos $(0, a, 0)$; y el eje z está formado por los puntos $(0, 0, a)$. También se suele denotar a los puntos en el espacio con las letras x, y y z en lugar de a, b y c. Así, la terna (x, y, z) representa un punto cuya primera coordenada es x, la segunda coordenada es y, y la tercera coordenada es z.

Empleamos la notación siguiente para la recta, el plano y el espacio tridimensional.

(i) La recta real se denota por $\mathbf{R}^1$ (así, es lo mismo $\mathbf{R}$ que $\mathbf{R}^1$).
(ii) El conjunto de todos los pares ordenados (x, y) de números reales se denota por $\mathbf{R}^2$.
(iii) El conjunto de todas las ternas ordenadas (x, y, z) de números reales se denota por $\mathbf{R}^3$.

Cuando se habla en conjunto de $\mathbf{R}^1$, $\mathbf{R}^2$ y $\mathbf{R}^3$, se escribe $\mathbf{R}^n$, $n = 1, 2$ o 3; o $\mathbf{R}^m$, $m = 1, 2, 3$.

La operación de suma se puede extender de $\mathbf{R}$ a $\mathbf{R}^2$ y $\mathbf{R}^3$. Para $\mathbf{R}^3$ se procede de la manera siguiente. Dadas dos ternas (x, y, z) y (x', y', z'), definimos su *suma* mediante

$$(x, y, z) + (x', y', z') = (x + x', y + y', z + z').$$

EJEMPLO 1

$$(1,1,1)+(2,-3,4)=(3,-2,5)$$
$$(x,y,z)+(0,0,0)=(x,y,z)$$
$$(1,7,3)+(2,0,6)=(3,7,9). \quad \blacktriangle$$

El elemento $(0,0,0)$ se llama *elemento cero* (o sólo *cero*) de $\mathbf{R}^3$. El elemento $(-x,-y,-z)$ se llama *inverso aditivo* (o *negativo*) de (x,y,z), y se escribe $(x,y,z)-(x',y',z')$ en lugar de $(x,y,z)+(-x',-y',-z')$.

Hay operaciones de producto que son importantes en $\mathbf{R}^3$. Una de ellas, llamada *producto interno*, asigna un número real a cada pareja de elementos de $\mathbf{R}^3$. En la sección 1.2. estudiaremos con detalle el producto interno. Otra operación de producto para $\mathbf{R}^3$ se llama *producto por un escalar* (la palabra "escalar" es sinónimo de "número real"). Este producto combina escalares (números reales) y elementos de $\mathbf{R}^3$ (ternas ordenadas) para producir elementos de $\mathbf{R}^3$ de la manera siguiente: dado un escalar α y una terna (x,y,z), definimos el *múltiplo escalar* o *producto por un escalar* mediante

$$\alpha(x,y,z)=(\alpha x,\alpha y,\alpha z).$$

EJEMPLO 2

$$2(4,e,1)=(2\cdot 4,2\cdot e,2\cdot 1)=(8,2e,2)$$
$$6(1,1,1)=(6,6,6)$$
$$1(x,y,z)=(x,y,z)$$
$$0(x,y,z)=(0,0,0)$$
$$\begin{aligned}(\alpha+\beta)(x,y,z)&=((\alpha+\beta)x,(\alpha+\beta)y,(\alpha+\beta)z)\\&=(\alpha x+\beta x,\alpha y+\beta y,\alpha z+\beta z)\\&=\alpha(x,y,z)+\beta(x,y,z). \quad \blacktriangle\end{aligned}$$

Como consecuencia de las definiciones, la suma y el producto por un escalar para $\mathbf{R}^3$ satisfacen las siguientes identidades:

(i) $(\alpha\beta)(x,y,z)=\alpha[\beta(x,y,z)]$ (asociatividad)

(ii) $(\alpha+\beta)(x,y,z)=\alpha(x,y,z)+\beta(x,y,z)$

(iii) $\alpha[(x,y,z)+(x',y',z')]=\alpha(x,y,z)+\alpha(x',y',z')$ (distributividad)

(iv) $\alpha(0,0,0)=(0,0,0)$

(v) $0(x,y,z)=(0,0,0)$ (propiedades del elemento cero)

(vi) $1(x,y,z)=(x,y,z)$ (propiedad del elemento identidad)

Para $\mathbf{R}^2$ se define la suma de la misma manera que para $\mathbf{R}^3$, mediante

$$(x,y)+(x',y')=(x+x',y+y'),$$

y el producto por un escalar se define como

$$\alpha(x, y) = (\alpha x, \alpha y).$$

Volvamos a la geometría de nuestro modelo. Una de las herramientas más poderosas de las matemáticas y sus aplicaciones ha sido el concepto de *vector*. Se define (geométricamente) un vector como un segmento de recta dirigido que comienza en el origen, esto es, un segmento de recta con magnitud y dirección especificados, con punto inicial en el origen. ¿Han oído decir a los pilotos "Estamos en el radio vector de la pista de aterrizaje"? Se refieren al vector que da la dirección y la distancia a que se encuentra el aeroplano de la pista de aterrizaje. Es inútil señalar lo importantes que son en este caso la dirección y la distancia. La figura 1.1.4 muestra varios vectores. Así, los vectores se pueden concebir como flechas que comienzan en el origen. Generalmente se imprimen en letras negritas: $\mathbf{v}$.

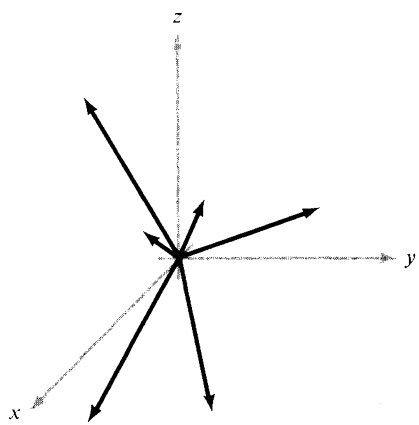


Figura 1.1.4 Los vectores se pueden concebir, geométricamente, como flechas saliendo del origen.

Usando esta definición de vector, podemos asociar con cada vector $\mathbf{v}$ el punto (x, y, z) en el espacio, donde termina $\mathbf{v}$, y, recíprocamente, a cada punto (x, y, z) en el espacio podemos asociar un vector $\mathbf{v}$. Así, identificaremos $\mathbf{v}$ con (x, y, z) y escribiremos $\mathbf{v} = (x, y, z)$. Por esta razón, los elementos de $\mathbf{R}^3$ no son sólo ternas ordenadas de números reales, sino que también se llaman vectores. La terna $(0, 0, 0)$ se denota por $\mathbf{0}$.

Decimos que dos vectores son ***iguales*** *si, y sólo si, tienen la misma dirección y la misma magnitud.* Esta condición se puede expresar de manera algebraica diciendo que si $\mathbf{v}_1 = (x, y, z)$ y $\mathbf{v}_2 = (x', y'z')$, entonces

$$\mathbf{v}_1 = \mathbf{v}_2 \quad \text{si, y sólo si,} \quad x = x',\ y = y',\ z = z'.$$

Geométricamente definimos el vector suma como sigue. En el plano que contiene a los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ (ver la figura 1.1.5), formemos el paralelogramo que

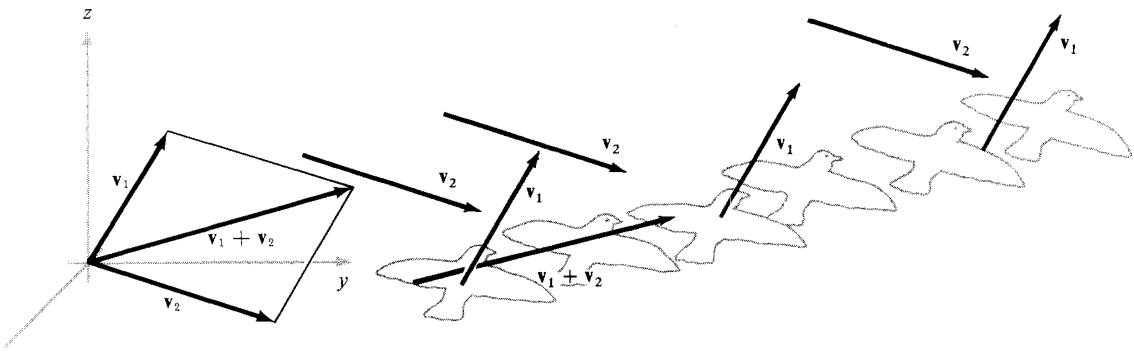


Figura 1.1.5 Geometría de la suma de vectores.

tiene como un lado a $\mathbf{v}_1$, y como lado adyacente a $\mathbf{v}_2$. Entonces la suma $\mathbf{v}_1 + \mathbf{v}_2$ es el segmento de recta dirigido a lo largo de la diagonal del paralelogramo. Esta consideración geométrica de la suma de vectores es útil en muchas situaciones físicas, como veremos más adelante. Para visualizar fácilmente esto mediante un ejemplo, consideren un ave o un aeroplano volando con velocidad $\mathbf{v}_1$, con un viento con velocidad $\mathbf{v}_2$. Lo que se ve es la velocidad resultante $\mathbf{v}_1 + \mathbf{v}_2$.

Para mostrar que la definición geométrica de la suma es consistente con la definición algebraica, debemos demostrar que $\mathbf{v}_1 + \mathbf{v}_2 = (x + x', y + y', z + z')$. Probaremos este resultado en el plano y dejaremos que el lector enuncie la proposición para el espacio tridimensional. Así, queremos mostrar que si $\mathbf{v}_1 = (x, y)$ y $\mathbf{v}_2 = (x', y')$, entonces $\mathbf{v}_1 + \mathbf{v}_2 = (x + x', y + y')$.

En la figura 1.1.6, sea $\mathbf{v}_1 = (x, y)$ el vector que termina en el punto A, y sea $\mathbf{v}_2 = (x', y')$ el vector que termina en el punto B. Por definición, el vector $\mathbf{v}_1 + \mathbf{v}_2$ termina en el vértice C del paralelogramo OBCA. Entonces, para verificar que

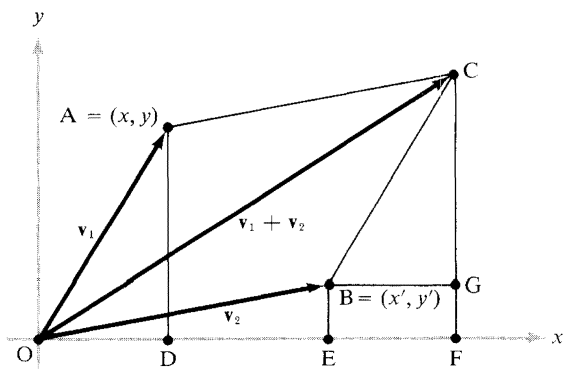


Figura 1.1.6 Construcción para la demostración de que $(x, y) + (x', y') = (x + x', y + y')$.

$\mathbf{v}_1 + \mathbf{v}_2 = (x + x', y + y')$, es suficiente mostrar que las coordenadas de C son $(x + x', y + y')$.

En la figura 1.1.6, los lados de los triángulos OAD y BCG son paralelos y los lados OA y BC tienen igual longitud, lo cual escribiremos como OA = BC. Por lo tanto, BG = OD; y como BGFE es un rectángulo, tenemos que EF = BG. Más aún, OD = x y OE = x'. De aquí que EF = BG = OD = x. Como OF = EF + OE, se sigue que OF = $x + x'$. Esto muestra que la coordenada x de C es $x+x'$. La demostración para la coordenada y es análoga. Con un argumento similar para los otros cuadrantes, vemos que la definición geométrica de la suma de vectores es equivalente a la definición algebraica en términos de coordenadas.

En la figura 1.1.7(a) se ilustra otra manera de considerar la suma vectorial: en términos de triángulos, en lugar de paralelogramos. Esto es, trasladamos (sin rotación) el segmento de recta dirigido que representa al vector $\mathbf{v}_2$, de modo que comience al final del vector $\mathbf{v}_1$. El punto final del segmento dirigido resultante es el punto final del vector $\mathbf{v}_1 + \mathbf{v}_2$. Notamos que cuando $\mathbf{v}_1$ y $\mathbf{v}_2$ son colineales, el triángulo se colapsa. Se ilustra esta situación en la figura 1.1.7(b).

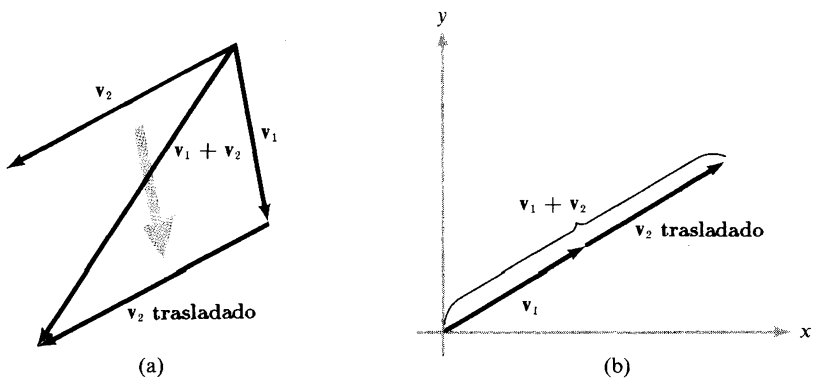


Figura 1.1.7 (a) Se puede visualizar la suma vectorial en términos de triángulos así como de paralelogramos. Sin embargo, el triángulo se colapsa cuando $\mathbf{v}_1$ y $\mathbf{v}_2$ son colineales (b).

Los múltiplos escalares de los vectores tienen interpretaciones geométricas similares. Si α es una escalar y $\mathbf{v}$ es un vector, definimos $\alpha\mathbf{v}$ como el vector que tiene α veces la longitud de $\mathbf{v}$, con la misma dirección que $\mathbf{v}$ si $\alpha > 0$, pero con dirección opuesta si $\alpha < 0$. La figura 1.1.8 ilustra varios ejemplos.

Al usar un razonamiento que depende de triángulos semejantes, podemos probar que si $\mathbf{v} = (x, y, z)$, entonces

$$\alpha\mathbf{v} = (\alpha x, \alpha y, \alpha z).$$

Esto es, la definición geométrica coincide con la algebraica.

¿Cómo representamos geométricamente al vector $\mathbf{b}-\mathbf{a}$? Como $\mathbf{a}+(\mathbf{b}-\mathbf{a}) = \mathbf{b}$, $\mathbf{b} - \mathbf{a}$ es el vector que al sumarlo a $\mathbf{a}$ da $\mathbf{b}$. En vista de esto, podemos concluir

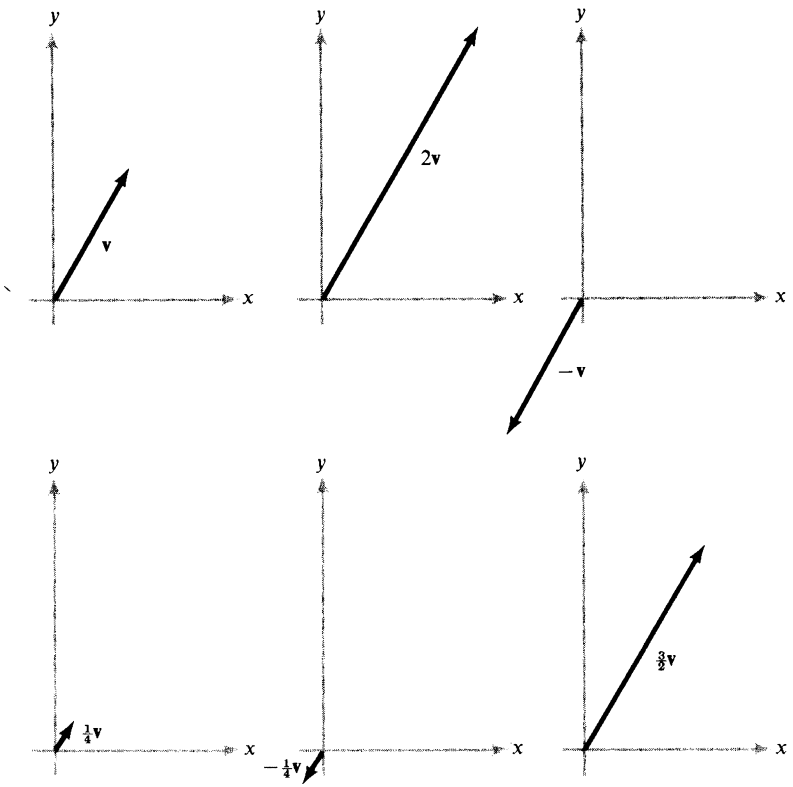


Figura 1.1.8 Algunos múltiplos escalares de un vector **v**.

que $\mathbf{b} - \mathbf{a}$ es el vector paralelo a, y con la misma magnitud que, el segmento de recta dirigido que comienza en el punto final de **a** y termina en el punto final de **b** (ver la figura 1.1.9).

Denotemos por **i** al vector que termina en $(1, 0, 0)$, por **j** al vector que termina en $(0, 1, 0)$ y por **k** al vector que termina en $(0, 0, 1)$. Por la definición de suma

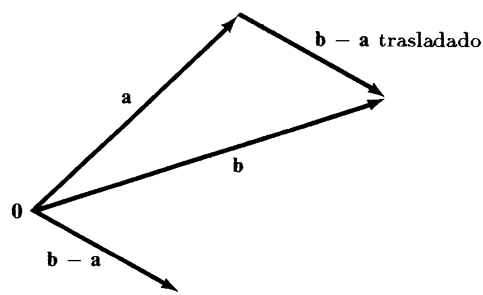


Figura 1.1.9 Geometría de la resta vectorial.

vectorial y la multiplicación por un escalar, hallamos que si $\mathbf{v} = (x, y, z)$, entonces

$$\mathbf{v} = x(1,0,0) + y(0,1,0) + z(0,0,1) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

Por lo tanto, podemos representar cualquier vector en el espacio tridimensional en términos de los vectores $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$. Es por esto que a los vectores $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$ se les llama *vectores de la base canónica* para $\mathbf{R}^3$.

EJEMPLO 3 El vector que termina en $(2,3,2)$ es $2\mathbf{i} + 3\mathbf{j} + 2\mathbf{k}$, y el vector que termina en $(0,-1,4)$ es $-\mathbf{j} + 4\mathbf{k}$. La figura 1.1.10 muestra a $2\mathbf{i} + 3\mathbf{j} + 2\mathbf{k}$; el lector deberá trazar el vector $-\mathbf{j} + 4\mathbf{k}$. ▲

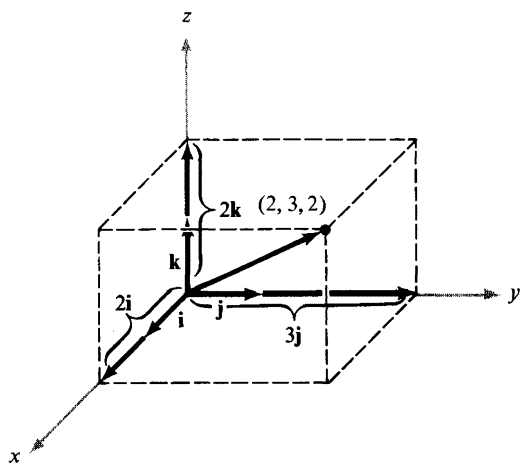


Figura 1.1.10 Representación de $(2,3,2)$ en términos de los vectores de la base canónica, $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$.

La suma y la multiplicación por un escalar se pueden escribir en términos de los vectores de la base canónica como sigue:

$$(x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) + (x'\mathbf{i} + y'\mathbf{j} + z'\mathbf{k}) = (x + x')\mathbf{i} + (y + y')\mathbf{j} + (z + z')\mathbf{k}$$

y

$$\alpha(x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = (\alpha x)\mathbf{i} + (\alpha y)\mathbf{j} + (\alpha z)\mathbf{k}.$$

Debido a la correspondencia entre puntos y vectores, a veces nos referimos al *punto* $\mathbf{a}$ en circunstancias en que se definió $\mathbf{a}$ como vector. El lector sobreentenderá que nos referimos al *punto final* del vector $\mathbf{a}$.

EJEMPLO 4 *Describir los puntos que están dentro del paralelogramo cuyos lados adyacentes son los vectores* **a** *y* **b**.

SOLUCIÓN Considerar la figura 1.1.11. Supongamos que P es cualquier punto dentro del paralelogramo dado y construimos las rectas l_1 y l_2 que pasan por P y son paralelas a los vectores **a** y **b**, respectivamente; vemos que l_1 interseca el lado del paralelogramo determinado por el vector **b** en algún punto $t\mathbf{b}$, donde $0 \leq t \leq 1$. Asimismo, l_2 interseca al lado determinado por el vector **a** en algún punto $s\mathbf{a}$, donde $0 \leq s \leq 1$.

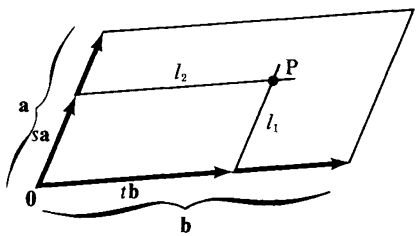


Figura 1.1.11 Descripción de los puntos dentro del paralelogramo formado por los vectores **a** y **b**.

Notar que P es el punto final de la diagonal de un paralelogramo con lados adyacentes $s\mathbf{a}$ y $t\mathbf{b}$; por lo tanto, si **v** denota al vector que termina en P, vemos que $\mathbf{v} = s\mathbf{a} + t\mathbf{b}$. Así, todos los puntos en el paralelogramo dado son puntos finales de vectores de la forma $s\mathbf{a} + t\mathbf{b}$ para $0 \leq s \leq 1$ y $0 \leq t \leq 1$. Regresando sobre nuestros pasos vemos que todos los vectores de esta forma terminan dentro del paralelogramo. ▲

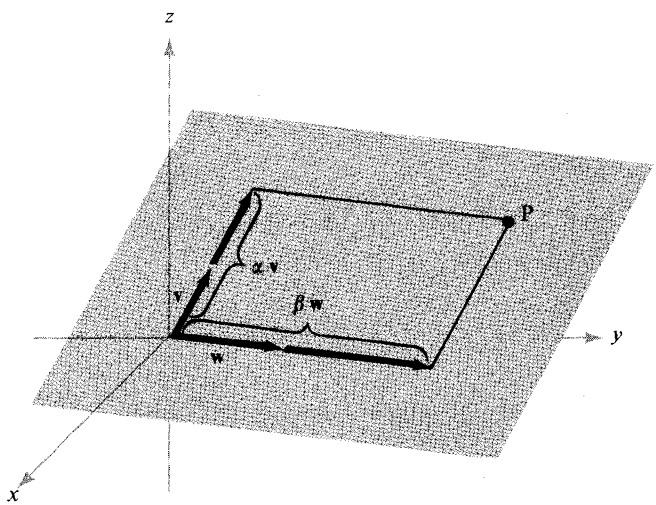


Figura 1.1.12 Descripción de los puntos P en el plano formado por los vectores **v** y **w**.

Como dos rectas que pasan por el origen determinan un plano que pasa por el origen, lo mismo sucede con dos vectores no paralelos. Si aplicamos el mismo razonamiento del ejemplo 4, vemos que el plano formado por dos vectores no paralelos $\mathbf{v}$ y $\mathbf{w}$ consta de todos los puntos de la forma $\alpha\mathbf{v} + \beta\mathbf{w}$, donde α y β varían sobre los números reales. Noten que cualquier punto P en el plano formado por los dos vectores será el vértice opuesto del paralelogramo determinado por $\alpha\mathbf{v}$ y $\beta\mathbf{w}$, donde α y β son algunos escalares, como en la figura 1.1.12.

El plano determinado por $\mathbf{v}$ y $\mathbf{w}$ se llama plano *generado por* $\mathbf{v}$ y $\mathbf{w}$. Cuando $\mathbf{v}$ es un múltiplo escalar de $\mathbf{w}$ y $\mathbf{w} \neq \mathbf{0}$, entonces $\mathbf{v}$ y $\mathbf{w}$ son paralelos y el plano degenera en una recta. Cuando $\mathbf{v} = \mathbf{w} = \mathbf{0}$ (esto es, cuando ambos son el vector cero), obtenemos un solo punto.

Hay tres planos particulares que surgen de manera natural en un sistema coordenado y que usaremos más adelante. Al plano generado por los vectores $\mathbf{i}$ y $\mathbf{j}$ se le llama plano xy, al plano generado por $\mathbf{j}$ y $\mathbf{k}$, plano yz, y al plano generado por $\mathbf{i}$ y $\mathbf{k}$, plano xz. Se ilustran estos planos en la figura 1.1.13.

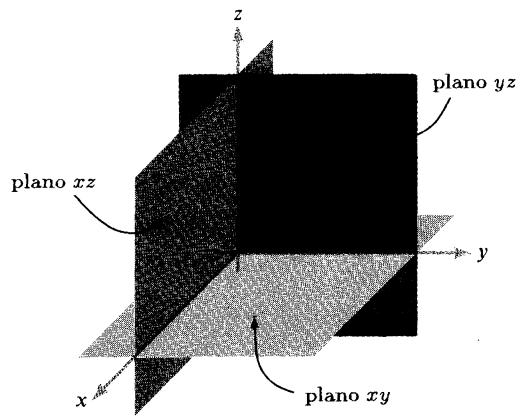


Figura 1.1.13 Los tres planos coordenados.

Los planos y las rectas son objetos geométricos que se pueden representar mediante ecuaciones. Pospondremos hasta la sección 1.3 el estudio de las ecuaciones que representan planos. Sin embargo, usando la interpretación geométrica de la suma vectorial y de la multiplicación por un escalar, podemos hallar la *ecuación de una recta l que pase por el punto final o extremo del vector* $\mathbf{a}$, *con la dirección de un vector* $\mathbf{v}$ (ver la figura 1.1.14). Conforme t varía por todos los números reales, los puntos de la forma $t\mathbf{v}$ son todos los múltiplos escalares del vector $\mathbf{v}$, y por lo tanto, agotan los puntos de la recta que pasa por el origen en la dirección de $\mathbf{v}$. Como todo punto sobre l es el extremo de la diagonal de un paralelogramo con lados $\mathbf{a}$ y $t\mathbf{v}$ para algún valor real de t, vemos que todos los puntos sobre l son de la forma $\mathbf{a} + t\mathbf{v}$. Así, la recta l se puede expresar mediante la ecuación $\mathbf{l}(t) = \mathbf{a} + t\mathbf{v}$. Decimos que l está expresada de manera *paramétrica*, con el parámetro t. En $t = 0$, $\mathbf{l}(t) = \mathbf{a}$. Cuando t crece, el punto $\mathbf{l}(t)$ se mueve

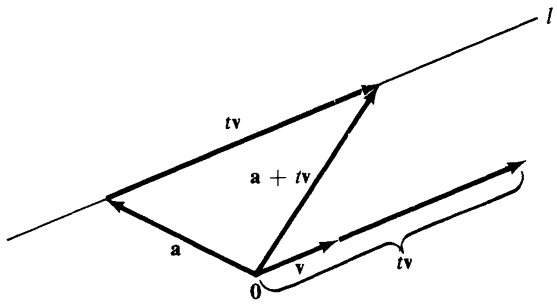


Figura 1.1.14 La recta l, dada en forma paramétrica por $\mathbf{l}(t) = \mathbf{a} + t\mathbf{v}$, está en dirección de $\mathbf{v}$ y pasa por la punta de $\mathbf{a}$.

alejándose de $\mathbf{a}$ en la dirección de $\mathbf{v}$. Conforme t decrece desde $t = 0$ por los valores negativos, $\mathbf{l}(t)$ se mueve alejándose de $\mathbf{a}$ en la dirección de $-\mathbf{v}$.

Puede haber varias *parametrizaciones* de la misma recta. Se pueden obtener escogiendo, en lugar de $\mathbf{a}$, un punto diferente sobre la recta dada, y formando la ecuación paramétrica de la recta comenzando en ese punto y en dirección de $\mathbf{v}$. Por ejemplo, el extremo de $\mathbf{a} + \mathbf{v}$ está sobre la recta $\mathbf{l}(t) = \mathbf{a} + t\mathbf{v}$, y así, $\mathbf{l}_1(t) = (\mathbf{a} + \mathbf{v}) + t\mathbf{v}$ representa la misma recta. Incluso se pueden obtener otras parametrizaciones observando que si $\alpha \neq 0$, el vector $\alpha\mathbf{v}$ tiene la misma dirección que $\mathbf{v}$ (o la opuesta). Así, $\mathbf{l}_2(t) = \mathbf{a} + t\alpha\mathbf{v}$ es otra parametrización de $\mathbf{l}(t) = \mathbf{a} + t\mathbf{v}$.

EJEMPLO 5 *Determinar la ecuación de la recta que pasa por* $(1, 0, 0)$ *en dirección de* $\mathbf{j}$.

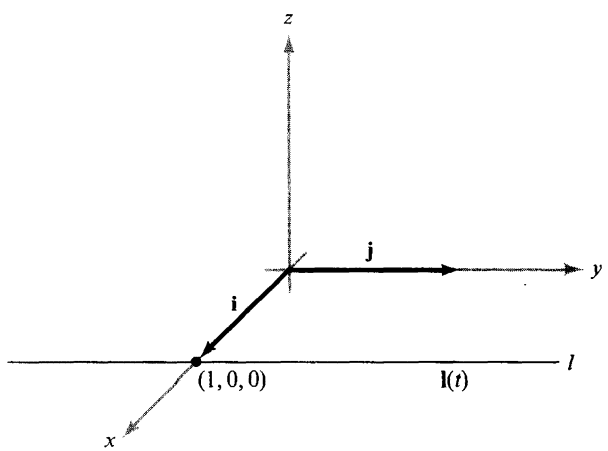


Figura 1.1.15 La recta l pasa por la punta de $\mathbf{i}$ en la dirección $\mathbf{j}$.

SOLUCIÓN La recta deseada se puede expresar en forma paramétrica como $\mathbf{l}(t) = \mathbf{i} + t\mathbf{j}$ (figura 1.1.15). En términos de coordenadas tenemos

$$\mathbf{l}(t) = (1,0,0) + t(0,1,0) = (1,t,0). \quad \blacktriangle$$

Vamos a deducir *la ecuación de una recta que pasa por los puntos finales de dos vectores dados* $\mathbf{a}$ *y* $\mathbf{b}$. Como el vector $\mathbf{b} - \mathbf{a}$ es paralelo al segmento de recta dirigido que va de $\mathbf{a}$ a $\mathbf{b}$, lo que deseamos es calcular la ecuación paramétrica de la recta que pasa por $\mathbf{a}$ en dirección de $\mathbf{b} - \mathbf{a}$ (figura 1.1.16). Así,

$$\mathbf{l}(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a}); \qquad \text{esto es,} \qquad \mathbf{l}(t) = (1-t)\mathbf{a} + t\mathbf{b}.$$

Conforme t crece de 0 a 1, sucede que $t(\mathbf{b} - \mathbf{a})$ comienza como el vector cero y crece en longitud (manteniéndose en la dirección de $\mathbf{b} - \mathbf{a}$) hasta que en $t = 1$ es el vector $\mathbf{b} - \mathbf{a}$. Así, para $\mathbf{l}(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a})$, conforme t crece de 0 a 1, el vector $\mathbf{l}(t)$ sé mueve de la punta de $\mathbf{a}$ a la punta de $\mathbf{b}$ a lo largo del segmento de recta dirigido de $\mathbf{a}$ a $\mathbf{b}$.

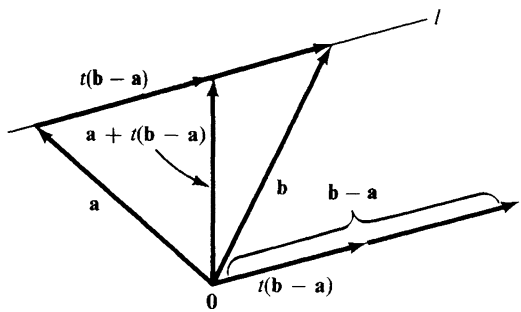


Figura 1.1.16 La recta l, dada en forma paramétrica por $\mathbf{l}(t) = \mathbf{a} + t(\mathbf{b} - \mathbf{a})$, pasa por las puntas de $\mathbf{a}$ y $\mathbf{b}$.

EJEMPLO 6 *Hallar la ecuación de la recta que pasa por* $(-1,1,0)$ *y* $(0,0,1)$ *(ver la figura 1.1.17).*

SOLUCIÓN Representemos los puntos dados por $\mathbf{a} = -\mathbf{i} + \mathbf{j}$ y $\mathbf{b} = \mathbf{k}$; tenemos

$$\begin{aligned}\mathbf{l}(t) &= (1-t)(-\mathbf{i} + \mathbf{j}) + t\mathbf{k} \\ &= -(1-t)\mathbf{i} + (1-t)\mathbf{j} + t\mathbf{k}.\end{aligned}$$

La ecuación de esta recta se puede escribir entonces como

$$\mathbf{l}(t) = (t-1)\mathbf{i} + (1-t)\mathbf{j} + t\mathbf{k},$$

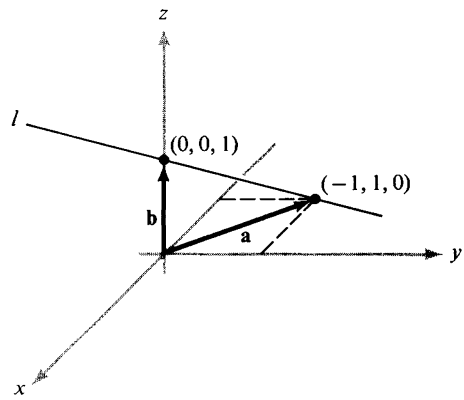


Figura 1.1.17 Caso especial de la figura anterior, donde $\mathbf{a} = (-1, 1, 0)$ y $\mathbf{b} = (0, 0, 1)$.

o, de manera equivalente, si $\mathbf{l}(t) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$,

$$x = t - 1, \qquad y = 1 - t, \qquad z = t. \quad \blacktriangle$$

En términos de componentes, *la ecuación de la recta que pasa por los dos puntos* (x_1, y_1, z_1) *y* (x_2, y_2, z_2) *es*

$$x = x_1 + t(x_2 - x_1), \qquad y = y_1 + t(y_2 - y_1), \qquad z = z_1 + t(z_2 - z_1)$$

Eliminando t es posible escribir esto como

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}.$$

Notamos que cualquier vector de la forma $\mathbf{c} = \lambda\mathbf{a} + \mu\mathbf{b}$, donde $\lambda + \mu = 1$, está sobre la *recta que pasa por los extremos de* $\mathbf{a}$ *y* $\mathbf{b}$. Para verlo, observar que $\mathbf{c} = (1 - \mu)\mathbf{a} + \mu\mathbf{b} = \mathbf{a} + \mu(\mathbf{b} - \mathbf{a})$.

EJEMPLO 7 *Usar métodos vectoriales para probar que las diagonales de un paralelogramo se bisecan entre sí.*

SOLUCIÓN Representemos los lados adyacentes del paralelogramo por los vectores $\mathbf{a}$ y $\mathbf{b}$, como se muestra en la figura 1.1.18. Primero calculamos el vector que va al punto medio del segmento de recta PQ. Como $\mathbf{b} - \mathbf{a}$ es paralelo e igual en longitud al segmento dirigido de P a Q, $(\mathbf{b} - \mathbf{a})/2$ es paralelo e igual en longitud al segmento de recta dirigido de P al punto medio de PQ. Así, el vector $\mathbf{a} + (\mathbf{b} - \mathbf{a})/2 = (\mathbf{a} + \mathbf{b})/2$ termina en el punto medio de PQ.

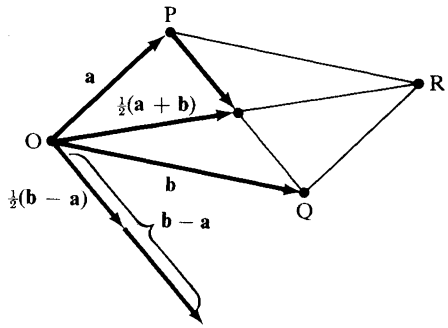


Figura 1.1.18 Construcciones usadas para demostrar que las diagonales de un paralelogramo se bisecan entre sí.

A continuación calculamos el vector que va al punto medio de OR. Sabemos que $\mathbf{a} + \mathbf{b}$ termina en R, de modo que $(\mathbf{a} + \mathbf{b})/2$ termina en el punto medio de OR. En vista de que ya probamos que el vector $(\mathbf{a} + \mathbf{b})/2$ termina en el punto medio de OR y en el punto medio de PQ, se sigue que OR y PQ se bisecan entre sí. ▲

Consideremos ahora algunas aplicaciones físicas de los vectores. Un ejemplo sencillo de cantidad física que se representa mediante un vector es un desplazamiento. Suponer que en una parte de la superficie terrestre lo suficientemente pequeña para considerarse plana, introducimos coordenadas de modo que el eje x apunte al este, el eje y apunte al norte, y la unidad de longitud sea el kilómetro. Si estamos en un punto P y queremos ir a un punto Q, el *vector de desplazamiento* $\mathbf{d}$ que une a P con Q nos indica la dirección y la distancia que tenemos que viajar. Si x y y son las componentes de este vector, el desplazamiento de P a Q es "x kilómetros al este, y kilómetros al norte".

EJEMPLO 8 *Supongan que dos navegantes que no se pueden ver entre sí, pero que se pueden comunicar por radio, quieren determinar la posición relativa de sus barcos. Explicar cómo pueden hacerlo si cada uno tiene la capacidad de determinar su vector de desplazamiento al mismo faro.*

SOLUCIÓN Sean P_1 y P_2 las posiciones de los barcos, y sea Q la posición del faro. El desplazamiento del i-ésimo barco al faro es el vector $\mathbf{d}_i$ que une a P_i con Q. El desplazamiento del primer barco al segundo es el vector $\mathbf{d}$ que une a P_1 con P_2. Tenemos que $\mathbf{d} + \mathbf{d}_2 = \mathbf{d}_1$ (figura 1.1.19), de modo que $\mathbf{d} = \mathbf{d}_1 - \mathbf{d}_2$. Esto es, el desplazamiento de un barco hasta el otro es la diferencia entre los desplazamientos desde los barcos hasta el faro. ▲

También podemos representar como vector la velocidad de un objeto en movimiento. Por el momento, sólo consideraremos objetos moviéndose con rapidez

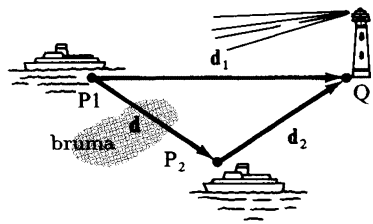


Figura 1.1.19 Se pueden usar métodos vectoriales para localizar objetos.

uniforme a lo largo de rectas. Supongan, por ejemplo, que un bote de vapor cruza un lago navegando a 10 kilómetros por hora (km/h) en dirección noreste. Después de 1 hora de viaje, el desplazamiento es $(10/\sqrt{2}, 10/\sqrt{2}) \approx (7.07, 7.07)$; ver la figura 1.1.20.

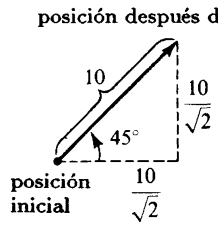


Figura 1.1.20 Si un objeto se mueve hacia el nordeste a 10 km/h, su vector velocidad tiene componentes $(10/\sqrt{2}, 10/\sqrt{2})$.

El vector cuyas componentes son $(10/\sqrt{2}, 10/\sqrt{2})$ se llama *vector velocidad* del bote. En general, si un objeto se mueve uniformemente a lo largo de una recta, *su vector velocidad es el vector desplazamiento desde la posición en cualquier momento hasta la posición en el momento* 1 *unidad de tiempo después.* Si aparece una corriente en el lago moviéndose hacia el este a 2 km/h, y el bote continúa apuntando hacia la misma dirección con el motor funcionando a la misma razón, su desplazamiento después de 1 hora tendrá las componentes dadas por

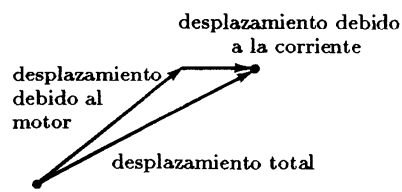


Figura 1.1.21 El desplazamiento total es la suma de los desplazamientos debidos al motor y a la corriente.

$(10/\sqrt{2}+2, 10/\sqrt{2})$; ver la figura 1.1.21. Por lo tanto el nuevo vector velocidad tiene componentes $(10/\sqrt{2}+2, 10/\sqrt{2})$. Notamos que ésta es la suma del vector velocidad original $(10/\sqrt{2}, 10/\sqrt{2})$ del bote y el vector velocidad $(2,0)$ de la corriente.

Si un objeto tiene vector velocidad (constante) $\mathbf{v}$, entonces en t segundos su vector desplazamiento resultante es $\mathbf{d} = t\mathbf{v}$; ver la figura 1.1.22.

Figura 1.1.22 Desplazamiento = tiempo × velocidad.

EJEMPLO 9 *Un ave va volando en línea recta con vector velocidad* $10\mathbf{i}+6\mathbf{j}+\mathbf{k}$ *(en kilómetros por hora). Suponer que* (x, y) *son sus coordenadas en tierra y que* z *es su altura.*

(a) *Si en cierto momento el ave está en la posición* $(1,2,3)$, *¿dónde estará una hora después? ¿Y un minuto después?*

(b) *¿Cuántos segundos tarda el ave en subir* 10 *metros?*

SOLUCIÓN (a) El vector desplazamiento desde $(1,2,3)$ después de 1 hora es $10\mathbf{i}+6\mathbf{j}+\mathbf{k}$, de modo que la nueva posición es $(1,2,3)+(10,6,1) = (11,8,4)$. Después de 1 minuto, el vector desplazamiento desde $(1,2,3)$ es $\frac{1}{60}(10\mathbf{i}+6\mathbf{j}+\mathbf{k})$ $= \frac{1}{6}\mathbf{i}+\frac{1}{10}\mathbf{j}+\frac{1}{60}\mathbf{k}$, de modo que la nueva posición es $(1,2,3)+(\frac{1}{6}, \frac{1}{10}, \frac{1}{60}) =$ $(\frac{7}{6}, \frac{21}{10}, \frac{181}{60})$.

(b) Después de t segundos $(= t/3600\,\text{h})$, el vector desplazamiento desde $(1,2,3)$ es $(t/3600)(10\mathbf{i}+6\mathbf{j}+\mathbf{k}) = (t/360)\mathbf{i}+(t/600)\mathbf{j}+(t/3600)\mathbf{k}$. El incremento en altura es la componente z $t/3600$. Esto es igual a 10 m $(=\frac{1}{100}\,\text{km})$ cuando $t/3600 = \frac{1}{100}$ —esto es, cuando $t = 36\,\text{s}$. ▲

EJEMPLO 10 *Las fuerzas físicas tienen magnitud y dirección, de modo que pueden representarse mediante vectores. Si actúan simultáneamente varias fuerzas sobre un objeto, la fuerza resultante está representada por la suma de los vectores de fuerza individuales. Suponer que las fuerzas* $\mathbf{i}+\mathbf{k}$ *y* $\mathbf{j}+\mathbf{k}$ *actúan sobre un cuerpo. ¿Qué tercera fuerza debemos imponer para contrarrestar a las dos —esto es, para hacer que la fuerza total sea igual a cero?*

SOLUCIÓN La fuerza $\mathbf{F}$ deberá escogerse de manera que $(\mathbf{i}+\mathbf{k})+(\mathbf{j}+\mathbf{k})+\mathbf{v} = \mathbf{0}$; esto es, $\mathbf{v} = -(\mathbf{i}+\mathbf{k})-(\mathbf{j}+\mathbf{k}) = -\mathbf{i}-\mathbf{j}-2\mathbf{k}$. (Recordar que $\mathbf{0}$ es el *vector cero*, el vector cuyas componentes son todas cero.) ▲

NOTA HISTÓRICA

Aproximadamente hasta el año de 1900 muchos científicos se resistieron a usar vectores, en favor de la teoría más complicada de los cuaterniones. El libro que popularizó los métodos vectoriales fue *Vector Analysis*, de E. B. Wilson (reimpreso por Dover en 1960), basado en los cursos impartidos por J. W. Gibbs en Yale en 1899 y 1900. Wilson se resistía a tomar el curso de Gibbs, pues había llevado en Harvard un curso de un año con J. M. Pierce, campeón en métodos con cuaterniones, pero un jefe de departamento lo obligó a añadir el curso a su programa. (Para más detalles ver *A History of Vector Analysis*, de M. J. Crowe, University of Notre Dame Press, Notre Dame, Ind., 1967.)

EJERCICIOS

(Los ejercicios que tienen números y letras dentro de un cuadro están resueltos en la Guía de estudio.)

Completar los cálculos en los ejercicios del 1 *al* 6.

1. $(-21, 23) - (?, 6) = (-25, ?)$

2. $3(133, -0.33, 0) + (-399, 0.99, 0) = (?, ?, ?)$

3. $(8a, -2b, 13c) = (52, 12, 11) + \frac{1}{2}(?, ?, ?)$

4. $(2, 3, 5) - 4\mathbf{i} + 3\mathbf{j} = (?, ?, ?)$

5. $800(0.03, 0, 0) = ?\mathbf{i} + ?\mathbf{j} + ?\mathbf{k}$

6. $(3, 4, 5) + (6, 2, -6) = (?, ?, ?)$

7. ¿Qué restricciones se deben tener sobre x, y y z de modo que la terna (x, y, z) represente un punto sobre el eje y? ¿Y sobre el eje z? ¿En el plano xz? ¿En el plano yz?

8. Trazar los vectores $\mathbf{v} = (2, 3, -6)$ y $\mathbf{w} = (-1, 1, 1)$. En esa figura, trazar $-\mathbf{v}$, $\mathbf{v}+\mathbf{w}$, $2\mathbf{v}$, y $\mathbf{v} - \mathbf{w}$.

9. (a) Generalizar la construcción geométrica en la figura 1.1.6 para mostrar que si $\mathbf{v}_1 = (x, y, z)$ y $\mathbf{v}_2 = (x', y', z')$ entonces $\mathbf{v}_1 + \mathbf{v}_2 = (x + x', y + y', z + z')$.

(b) Usando un argumento basado en triángulos semejantes, probar que $\alpha\mathbf{v} = (\alpha x, \alpha y, \alpha z)$ cuando $\mathbf{v} = (x, y, z)$.

10. Repetir el ejercicio 8 usando $\mathbf{v} = (2, 1, 3)$ y $\mathbf{w} = (-2, 0, -1)$.

En los ejercicios del 11 al 17, usar notación de conjuntos o vectorial, o ambas, para describir los puntos que están en las configuraciones dadas, como lo hicimos en los ejemplos 4, 5 y 6.

11. El plano generado por $\mathbf{v}_1 = (2, 7, 0)$ y $\mathbf{v}_2 = (0, 2, 7)$.

12. El plano generado por $\mathbf{v}_1 = (3, -1, 1)$ y $\mathbf{v}_2 = (0, 3, 4)$.

13. La recta que pasa por $(-1, -1, -1)$ en la dirección de $\mathbf{j}$.

14. La recta que pasa por $(0, 2, 1)$ en la dirección de $2\mathbf{i} - \mathbf{k}$.

15. La recta que pasa por $(-1, -1, -1)$ y $(1, -1, 2)$.

16. La recta que pasa por $(-5, 0, 4)$ y $(6, -3, 2)$.

17. El paralelogramo cuyos lados adyacentes son los vectores $\mathbf{i} + 3\mathbf{k}$ y $-2\mathbf{j}$.

18. Hallar los puntos de intersección de la recta $x = 3 + 2t$, $y = 7 + 8t$, $z = -2 + t$, esto es, $\mathbf{l}(t) = (3 + 2t, 7 + 8t, -2 + t)$, con los planos coordenados.

19. Mostrar que no hay puntos (x, y, z) que satisfagan $2x - 3y + z - 2 = 0$ y que estén sobre la recta $\mathbf{v} = (2, -2, -1) + t(1, 1, 1)$.

20. Mostrar que todo punto sobre la recta $\mathbf{v} = (1, -1, 2) + t(2, 3, 1)$ satisface $5x - 3y - z - 6 = 0$.

21. Mostrar que las medianas de un triángulo se intersecan en un punto, y que este punto divide a cada mediana con una razón $2 : 1$.

En los ejercicios del 22 al 24, usar métodos vectoriales para describir las configuraciones dadas.

22. El paralelepípedo que tiene como aristas a los vectores $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$. (La región que tenemos en mente está en la figura 1.3.5.)

23. Los puntos dentro del paralelogramo con una esquina en (x_0, y_o, z_0) tal que los lados que salen de esa esquina son iguales en magnitud y dirección a los vectores $\mathbf{a}$ y $\mathbf{b}$.

24. El plano determinado por los tres puntos (x_0, y_0, z_0), (x_1, y_1, z_1) y (x_2, y_2, z_2).

25. Un barco situado en la posición $(1, 0)$ en una carta de navegación (con el norte en la dirección y positiva) divisa una roca en la posición $(2, 4)$. ¿Cuál es el vector que une al barco con la roca? ¿Qué ángulo θ forma este vector con la dirección norte? (Se le llama la *orientación* de la roca desde el barco.)

26. Supongan que el barco del ejercicio 25 apunta al rumbo norte y viaja con una rapidez de 4 nudos respecto al agua. Hay una corriente que fluye con dirección este a 1 nudo; las unidades de la carta son millas náuticas; 1 nudo = 1 milla náutica por hora.

(a) Si no hubiera corriente, ¿qué vector $\mathbf{u}$ representaría la velocidad del barco respecto al fondo del mar?

(b) Si el barco siguiera la corriente, ¿qué vector $\mathbf{v}$ representaría su velocidad respecto al fondo del mar?

(c) ¿Qué vector $\mathbf{w}$ representa la velocidad total del barco?

(d) ¿Dónde estará el barco después de una hora?

(e) ¿Deberá cambiar el rumbo el capitán?

(f) ¿Qué pasaría si la roca fuera un iceberg?

27. Un aeroplano está situado en la posición $(3, 4, 5)$ al mediodía, y viaja con velocidad $400\mathbf{i} + 500\mathbf{j} - \mathbf{k}$ kilómetros por hora. El piloto sabe que hay un aeropuerto en la posición $(23, 29, 0)$.

(a) ¿A qué hora pasará el avión directamente sobre el aeropuerto? (Suponer que la Tierra es plana y que el vector $\mathbf{k}$ apunta hacia arriba.)

(b) ¿Cuál será la altura del avión cuando pase?

28. La velocidad $\mathbf{v}_1$ del viento es de 40 millas por hora (mi/h) de este a oeste, mientras que un aeroplano viaja con velocidad en el aire $\mathbf{v}_2$ de 100 mi/h con rumbo norte. La rapidez del aeroplano respecto a la Tierra es el vector suma $\mathbf{v}_1 + \mathbf{v}_2$.

(a) Hallar $\mathbf{v}_1 + \mathbf{v}_2$.

(b) Trazar una figura a escala.

29. Una fuerza de 50 lb se dirige a 50° sobre la horizontal, apuntando a la derecha. Determinar sus componentes horizontal y vertical. Mostrar los resultados en una figura.

30. Dos personas jalan horizontalmente de cuerdas atadas a un poste; el ángulo entre las cuerdas es de 60°. A jala con una fuerza de 150 lb, mientras que B jala con una fuerza de 110 lb.

(a) La fuerza resultante es la suma vectorial de las dos fuerzas en un sistema coordenado escogido de manera conveniente. Trazar una figura a escala que represente gráficamente a las tres fuerzas.

(b) Usando trigonometría, determinar fórmulas para las componentes vectoriales de las dos fuerzas en un sistema coordenado escogido de manera conveniente. Efectuar la suma algebraica y hallar el ángulo que la fuerza resultante hace con A.

31. 1 kilogramo (kg) masa situado en el origen se cuelga de cuerdas fijadas en los puntos $(1, 1, 1)$ y $(-1, -1, 1)$. Si la fuerza de gravedad apunta en la direccción del vector $-\mathbf{k}$, ¿cuál es el vector que describe la fuerza a lo largo de cada cuerda? [IDEA: Usar la simetría del problema. 1 kg masa pesa 9.8 newtons (N).]

32. Escribir la ecuación química $CO + H_2O = H_2 + CO_2$ como una ecuación en ternas ordenadas (C, O, H), e ilustrarla mediante un diagrama vectorial en el espacio.

33. (a) Escribir la ecuación química $pC_3H_4O_3 + qO_2 = rCO_2 + sH_2O$ como una ecuación en ternas ordenadas con coeficientes desconocidos p, q, r y s.

(b) Hallar la menor solución entera positiva posible para p, q, r y s,

(c) Ilustrar la solución mediante un diagrama vectorial en el espacio.

***34.** Hallar una recta que esté en el conjunto definido por la ecuación $x^2 + y^2 - z^2 = 1$.

1.2 EL PRODUCTO INTERNO

En ésta y en la sección siguiente estudiaremos dos productos de vectores: el producto interno y el producto cruz. Son muy útiles en aplicaciones físicas y tienen interpretaciones geométricas interesantes. El primer producto que vamos a considerar se llama *producto interno*. Con frecuencia se le llama también *producto punto*.

Supongamos que tenemos dos vectores $\mathbf{a}$ y $\mathbf{b}$ en $\mathbf{R}^3$ (figura 1.2.1) y queremos determinar el ángulo entre ellos, esto es, el menor ángulo subtendido por $\mathbf{a}$ y $\mathbf{b}$ en el plano que generan. El producto interno nos permite hacerlo. Primero desarrollamos formalmente el concepto y después probamos que este producto hace lo que aseguramos. Sea $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ y $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$. Definimos el *producto interno* de $\mathbf{a}$ y $\mathbf{b}$, que se escribe como $\mathbf{a} \cdot \mathbf{b}$, como el número real

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3.$$

Noten que el producto interno de dos vectores es una cantidad escalar. A veces se denota al producto interno por $\langle \mathbf{a}, \mathbf{b} \rangle$. Es frecuente hacerlo por razones tipográficas. *Así,* $\langle \mathbf{a}, \mathbf{b} \rangle$ *y* $\mathbf{a} \cdot \mathbf{b}$ *significan exactamente lo mismo.*

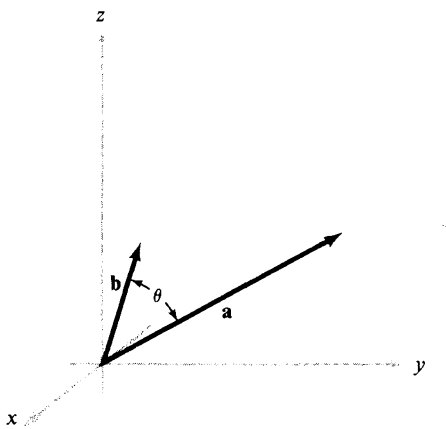


Figura 1.2.1 θ es el ángulo entre los vectores $\mathbf{a}$ y $\mathbf{b}$.

A partir de la definición se siguen ciertas propiedades del producto interno. Si $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ son vectores en $\mathbf{R}^3$ y α y β son números reales, entonces

(i) $\mathbf{a} \cdot \mathbf{a} \geq 0$;
$\mathbf{a} \cdot \mathbf{a} = 0$ si, y sólo si, $\mathbf{a} = \mathbf{0}$.

(ii) $\alpha\mathbf{a} \cdot \mathbf{b} = \alpha(\mathbf{a} \cdot \mathbf{b})$ y $\mathbf{a} \cdot \beta\mathbf{b} = \beta(\mathbf{a} \cdot \mathbf{b})$.

(iii) $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$ y $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$.

(iv) $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$.

Para probar la primera de estas propiedades, observen que si $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, entonces $\mathbf{a}\cdot\mathbf{a} = a_1^2 + a_2^2 + a_3^2$. Como a_1, a_2 y a_3 son números reales, sabemos que $a_1^2 \geq 0$, $a_2^2 \geq 0$, $a_3^2 \geq 0$. Así, $\mathbf{a}\cdot\mathbf{a} \geq 0$. Más aún, si $a_1^2 + a_2^2 + a_3^2 = 0$, entonces $a_1 = a_2 = a_3 = 0$, por lo tanto $\mathbf{a} = \mathbf{0}$ (el vector cero). Las demostraciones de las otras propiedades del producto interno también se obtienen fácilmente.

Se sigue del teorema de Pitágoras que la *longitud* del vector $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ es $\sqrt{a_1^2 + a_2^2 + a_3^2}$ (ver la figura 1.2.2). La longitud del vector $\mathbf{a}$ se denota por $\|\mathbf{a}\|$. Es frecuente llamar a esta cantidad la *norma* de $\mathbf{a}$. Como $\mathbf{a}\cdot\mathbf{a} = a_1^2 + a_2^2 + a_3^2$, se sigue que

$$\|\mathbf{a}\| = (\mathbf{a}\cdot\mathbf{a})^{1/2}.$$

Los vectores que tienen norma 1 se llaman *vectores unitarios*. Por ejemplo, los vectores $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$ son vectores unitarios. Observar que para cualquier vector distinto de cero $\mathbf{a}$, $\mathbf{a}/\|\mathbf{a}\|$ es un vector unitario; cuando dividimos $\mathbf{a}$ entre $\|\mathbf{a}\|$, decimos que hemos *normalizado* $\mathbf{a}$.

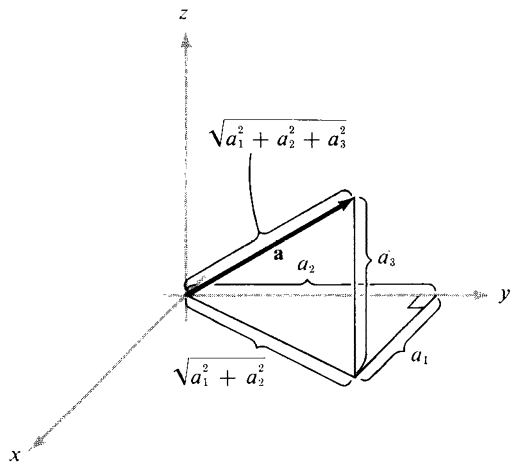


Figura 1.2.2 La longitud del vector $\mathbf{a} = (a_1, a_2, a_3)$ está dada por la fórmula pitagórica: $\sqrt{a_1^2 + a_2^2 + a_3^2}$.

En el plano, definir el vector $\mathbf{i}_\theta = (\cos\theta)\mathbf{i} + (\operatorname{sen}\theta)\mathbf{j}$, que es el vector unitario que forma un ángulo θ con el eje x (ver la figura 1.2.3). Claramente,

$$\|\mathbf{i}_\theta\| = (\operatorname{sen}^2\theta + \cos^2\theta)^{1/2} = 1.$$

Si $\mathbf{a}$ y $\mathbf{b}$ son vectores, hemos visto que el vector $\mathbf{b} - \mathbf{a}$ es paralelo a, y tiene la misma magnitud que el segmento de recta dirigido que va del extremo de $\mathbf{a}$ al extremo de $\mathbf{b}$. Se sigue que la distancia del extremo de $\mathbf{a}$ al extremo de $\mathbf{b}$ es $\|\mathbf{b} - \mathbf{a}\|$ (ver la figura 1.2.4).

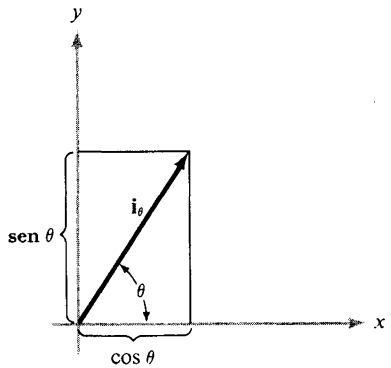


Figura 1.2.3 Las coordenadas de $\mathbf{i}_\theta$ son $\cos\theta$ y $\operatorname{sen}\theta$.

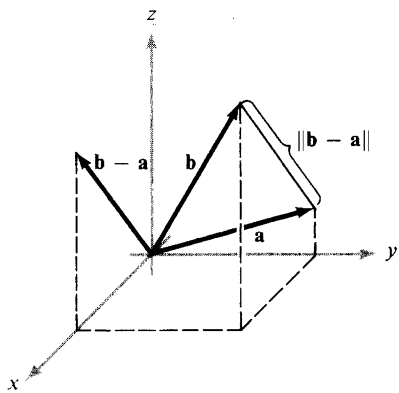


Figura 1.2.4 La distancia entre las puntas de **a** y **b** es $\|\mathbf{b}-\mathbf{a}\|$.

EJEMPLO 1 *Hallar la distancia del extremo del vector* **i**, *esto es, del punto* $(1,0,0)$ *al extremo del vector* **j**, $(0,1,0)$.

SOLUCIÓN $\|\mathbf{j}-\mathbf{i}\| = \sqrt{(0-1)^2+(1-0)^2+(0-0)^2} = \sqrt{2}.$ ▲

Mostremos ahora que el producto interno en efecto mide el ángulo entre dos vectores.

TEOREMA 1 *Sean* **a** *y* **b** *dos vectores en* $\mathbf{R}^3$ *y sea* θ, $0 \leq \theta \leq \pi$, *el ángulo entre ellos (figura 1.2.5). Entonces*

$$\mathbf{a}\cdot\mathbf{b} = \|\mathbf{a}\|\,\|\mathbf{b}\|\cos\theta.$$

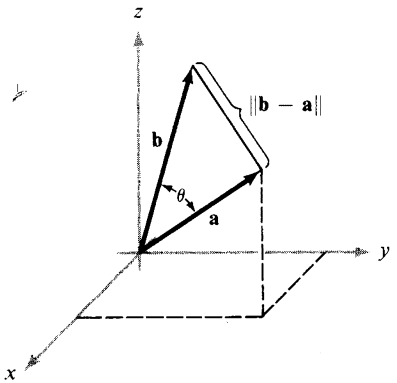


Figura 1.2.5 Los vectores **a**, **b** y el ángulo θ entre ellos; geometría del teorema 1 y su demostración.

De modo que podemos expresar el ángulo entre **a** *y* **b** *como*

$$\theta = \cos^{-1}\left(\frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\,\|\mathbf{b}\|}\right)$$

si **a** *y* **b** *son vectores distintos de cero.*

DEMOSTRACIÓN Si aplicamos la ley de los cosenos, aprendida en trigonometría, al triángulo con un vértice en el origen y lados adyacentes determinados por los vectores **a** y **b**, se sigue que

$$\|\mathbf{b}-\mathbf{a}\|^2 = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - 2\|\mathbf{a}\|\,\|\mathbf{b}\|\cos\theta.$$

Como $\|\mathbf{b}-\mathbf{a}\|^2 = (\mathbf{b}-\mathbf{a})\cdot(\mathbf{b}-\mathbf{a})$, $\|\mathbf{a}\|^2 = \mathbf{a}\cdot\mathbf{a}$, y $\|\mathbf{b}\|^2 = \mathbf{b}\cdot\mathbf{b}$, podemos reescribir la ecuación anterior como

$$(\mathbf{b}-\mathbf{a})\cdot(\mathbf{b}-\mathbf{a}) = \mathbf{a}\cdot\mathbf{a} + \mathbf{b}\cdot\mathbf{b} - 2\|\mathbf{a}\|\,\|\mathbf{b}\|\cos\theta.$$

Ahora,

$$\begin{aligned}(\mathbf{b}-\mathbf{a})\cdot(\mathbf{b}-\mathbf{a}) &= \mathbf{b}\cdot(\mathbf{b}-\mathbf{a}) - \mathbf{a}\cdot(\mathbf{b}-\mathbf{a})\\ &= \mathbf{b}\cdot\mathbf{b} - \mathbf{b}\cdot\mathbf{a} - \mathbf{a}\cdot\mathbf{b} + \mathbf{a}\cdot\mathbf{a}\\ &= \mathbf{a}\cdot\mathbf{a} + \mathbf{b}\cdot\mathbf{b} - 2\mathbf{a}\cdot\mathbf{b}.\end{aligned}$$

Así,

$$\mathbf{a}\cdot\mathbf{a} + \mathbf{b}\cdot\mathbf{b} - 2\mathbf{a}\cdot\mathbf{b} = \mathbf{a}\cdot\mathbf{a} + \mathbf{b}\cdot\mathbf{b} - 2\|\mathbf{a}\|\,\|\mathbf{b}\|\cos\theta.$$

Esto es,

$$\mathbf{a}\cdot\mathbf{b} = \|\mathbf{a}\|\,\|\mathbf{b}\|\cos\theta. \qquad \blacksquare$$

Este resultado muestra que el producto interno de dos vectores es el producto de sus longitudes por el coseno del ángulo entre ellos. Esta relación es útil con frecuencia en problemas de naturaleza geométrica.

COROLARIO (DESIGUALDAD DE CAUCHY-SCHWARZ) *Para cualesquiera dos vectores* $\mathbf{a}$ *y* $\mathbf{b}$, *tenemos*

$$|\mathbf{a} \cdot \mathbf{b}| \leq \|\mathbf{a}\| \, \|\mathbf{b}\|$$

con la igualdad si y sólo si $\mathbf{a}$ *es un múltiplo escalar de* $\mathbf{b}$, *o uno de ellos es* $\mathbf{0}$.

DEMOSTRACIÓN Si $\mathbf{a}$ no es un múltiplo escalar de $\mathbf{b}$, entonces $|\cos\theta| < 1$ y se cumple la desigualdad. Cuando $\mathbf{a}$ es un múltiplo escalar de $\mathbf{b}$, entonces $\theta = 0$ o π y $|\cos\theta| = 1$. ■

EJEMPLO 2 *Hallar el ángulo entre los vectores* $\mathbf{i}+\mathbf{j}+\mathbf{k}$ *e* $\mathbf{i}+\mathbf{j}-\mathbf{k}$ *(figura 1.2.6).*

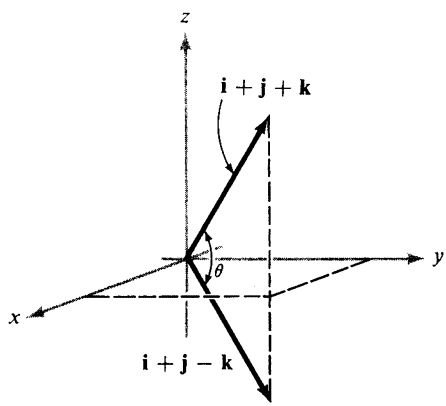


Figura 1.2.6 Búsqueda del ángulo entre $\mathbf{a} = \mathbf{i}+\mathbf{j}+\mathbf{k}$ y $\mathbf{b} = \mathbf{i}+\mathbf{j}-\mathbf{k}$.

SOLUCIÓN Usando el teorema 1, tenemos

$$(\mathbf{i}+\mathbf{j}+\mathbf{k}) \cdot (\mathbf{i}+\mathbf{j}-\mathbf{k}) = \|\mathbf{i}+\mathbf{j}+\mathbf{k}\| \, \|\mathbf{i}+\mathbf{j}-\mathbf{k}\| \cos\theta$$

de modo que

$$1 + 1 - 1 = (\sqrt{3})(\sqrt{3}) \cos\theta.$$

De donde

$$\cos\theta = \tfrac{1}{3}.$$

Esto es,

$$\theta = \cos^{-1}(\tfrac{1}{3}) \approx 1.23 \text{ radianes } (71°). \quad \blacktriangle$$

Si $\mathbf{a}$ y $\mathbf{b}$ son vectores distintos de cero en $\mathbf{R}^3$ y θ es el ángulo entre ellos, vemos que $\mathbf{a} \cdot \mathbf{b} = 0$ si y sólo si $\cos\theta = 0$. Así, *el producto interno de dos vectores distintos de cero es cero si y sólo si los vectores son perpendiculares.* Por lo tanto el producto interno nos proporciona un buen método para determinar si dos vectores son perpendiculares. Se suele decir que los vectores perpendiculares son *ortogonales.* Los vectores de la base canónica, $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$ son ortogonales entre sí, y tienen longitud 1; dichos sistemas se llaman *ortonormales.* Adoptaremos la convención de que el vector cero es ortogonal a todos los vectores.

EJEMPLO 3 *Los vectores* $\mathbf{i}_\theta = (\cos\theta)\mathbf{i} + (\operatorname{sen}\theta)\mathbf{j}$ *y* $j_\theta = -(\operatorname{sen}\theta)\mathbf{i} + (\cos\theta)\mathbf{j}$ *son ortogonales, pues*

$$\mathbf{i}_\theta \cdot \mathbf{j}_\theta = -\cos\theta \operatorname{sen}\theta + \operatorname{sen}\theta \cos\theta = 0$$

(ver la figura 1.2.7). $\blacktriangle$

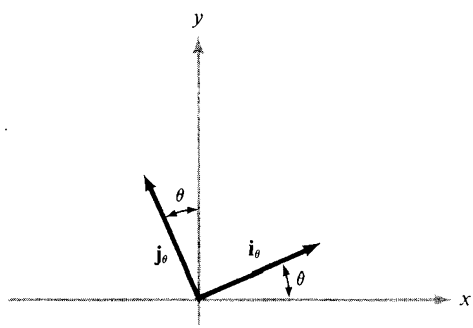


Figura 1.2.7 Los vectores $\mathbf{i}_\theta$ y $\mathbf{j}_\theta$ son ortogonales.

EJEMPLO 4 *Sean* $\mathbf{a}$ *y* $\mathbf{b}$ *dos vectores ortogonales distintos de cero. Si* $\mathbf{c}$ *es un vector en el plano generado por* $\mathbf{a}$ *y* $\mathbf{b}$, *entonces hay escalares* α *y* β *tales que* $\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b}$. *Usar el producto interno para determinar* α *y* β *(ver la figura 1.2.8).*

SOLUCIÓN Tomando el producto interno de $\mathbf{a}$ y $\mathbf{c}$, tenemos

$$\mathbf{a} \cdot \mathbf{c} = \mathbf{a} \cdot (\alpha\mathbf{a} + \beta\mathbf{b}) = \alpha\mathbf{a} \cdot \mathbf{a} + \beta\mathbf{a} \cdot \mathbf{b}.$$

Como $\mathbf{a}$ y $\mathbf{b}$ son ortogonales, $\mathbf{a} \cdot \mathbf{b} = 0$, de modo que,

$$\alpha = \frac{\mathbf{a} \cdot \mathbf{c}}{\mathbf{a} \cdot \mathbf{a}} = \frac{\mathbf{a} \cdot \mathbf{c}}{\|\mathbf{a}\|^2}.$$

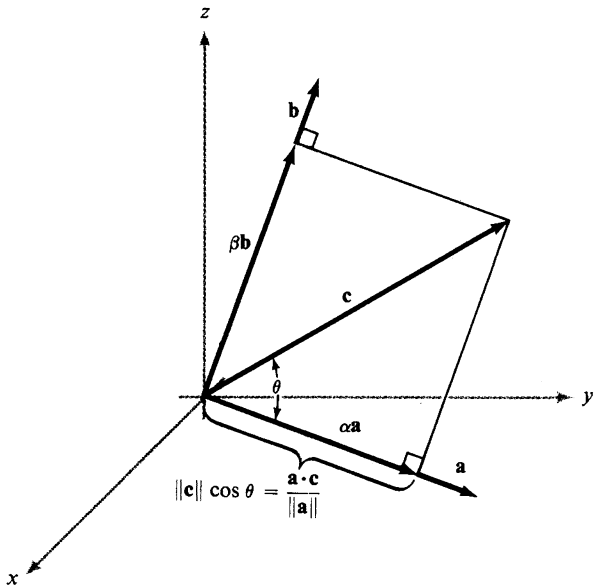


Figura 1.2.8 La geometría para la búsqueda de α y β donde $\mathbf{c} = \alpha\mathbf{a} + \beta\mathbf{b}$, como en el ejemplo 4.

De manera análoga,

$$\beta = \frac{\mathbf{b}\cdot\mathbf{c}}{\mathbf{b}\cdot\mathbf{b}} = \frac{\mathbf{b}\cdot\mathbf{c}}{\|\mathbf{b}\|^2}. \quad \blacktriangle$$

En este ejemplo, el vector $\alpha\mathbf{a}$ se llama la *proyección* de $\mathbf{c}$ a lo largo de $\mathbf{a}$, y $\beta\mathbf{b}$ es su *proyección* a lo largo de $\mathbf{b}$.

El resultado del ejemplo 4 también se puede obtener usando la interpretación geométrica del producto interno. Sea l la distancia medida a lo largo de la recta determinada al extender $\mathbf{a}$, del origen al punto donde la perpendicular desde $\mathbf{c}$ interseca a la extensión de $\mathbf{a}$. Se sigue que

$$l = \|\mathbf{c}\| \cos\theta$$

donde θ es el ángulo entre $\mathbf{a}$ y $\mathbf{c}$. Más aún, $l = \alpha\|\mathbf{a}\|$. Juntando estos resultados tenemos

$$\alpha\|\mathbf{a}\| = \|\mathbf{c}\| \cos\theta, \qquad \text{o} \qquad \alpha = \frac{\|\mathbf{c}\| \cos\theta}{\|\mathbf{a}\|} = \frac{\|\mathbf{c}\|}{\|\mathbf{a}\|}\left(\frac{\mathbf{a}\cdot\mathbf{c}}{\|\mathbf{c}\|\,\|\mathbf{a}\|}\right) = \frac{\mathbf{a}\cdot\mathbf{c}}{\mathbf{a}\cdot\mathbf{a}}.$$

Así, *la proyección de* $\mathbf{c}$ *sobre* $\mathbf{a}$ *está dada por*

$$\frac{\mathbf{c}\cdot\mathbf{a}}{\mathbf{a}\cdot\mathbf{a}}\mathbf{a} = \frac{\mathbf{c}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a}.$$

Notar que la longitud de la proyección de un vector $\mathbf{c}$ sobre un vector $\mathbf{a}$, donde θ es el ángulo entre $\mathbf{a}$ y $\mathbf{c}$, está dada por

$$\|\mathbf{c}\|\,|\cos\theta| = \frac{|\mathbf{a}\cdot\mathbf{c}|}{\|\mathbf{a}\|}$$

(ver la figura 1.2.9).

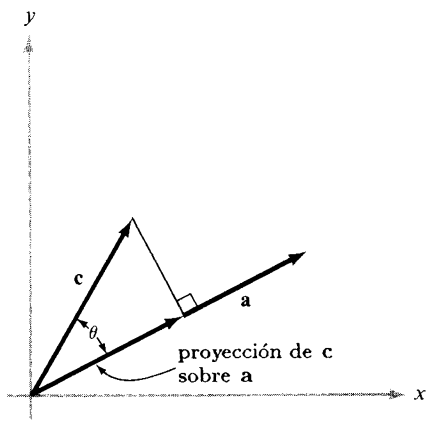


Figura 1.2.9 La proyección de $\mathbf{c}$ sobre $\mathbf{a}$ es $(\mathbf{a}\cdot\mathbf{c}/\|\mathbf{a}\|^2)\mathbf{a}$.

EJERCICIOS

1. (a) Probar las propiedades (ii) y (iii) del producto interno.
(b) Probar que $\mathbf{a}\cdot\mathbf{b} = \mathbf{b}\cdot\mathbf{a}$.

2. Calcular $\mathbf{a}\cdot\mathbf{b}$ donde $\mathbf{a} = 2\mathbf{i} + 10\mathbf{j} - 12\mathbf{k}$ y $\mathbf{b} = -3\mathbf{i} + 4\mathbf{k}$.

3. Hallar el ángulo entre $7\mathbf{j} + 19\mathbf{k}$ y $-2\mathbf{i} - \mathbf{j}$ (al grado más cercano).

4. Calcular $\mathbf{u}\cdot\mathbf{v}$, donde $\mathbf{u} = \sqrt{3}\mathbf{i} - 315\mathbf{j} + 22\mathbf{k}$ y $\mathbf{v} = \mathbf{u}/\|\mathbf{u}\|$.

5. ¿Es igual a cero $\|8\mathbf{i} - 12\mathbf{k}\|\cdot\|6\mathbf{j} + \mathbf{k}\| - |(8\mathbf{i} - 12\mathbf{k})\cdot(6\mathbf{j} + \mathbf{k})|$? Explicar.

En los ejercicios del 6 al 11, calcular $\|\mathbf{u}\|$, $\|\mathbf{v}\|$ *y* $\mathbf{u}\cdot\mathbf{v}$ *para los vectores dados en* $\mathbf{R}^3$.

6. $\mathbf{u} = 15\mathbf{i} - 2\mathbf{j} + 4\mathbf{k}$, $\mathbf{v} = \pi\mathbf{i} + 3\mathbf{j} - \mathbf{k}$

7. $\mathbf{u} = 2\mathbf{j} - \mathbf{i}$, $\mathbf{v} = -\mathbf{j} + \mathbf{i}$

8. $\mathbf{u} = 5\mathbf{i} - \mathbf{j} + 2\mathbf{k}$, $\mathbf{v} = \mathbf{i} + \mathbf{j} - \mathbf{k}$

9. $\mathbf{u} = -\mathbf{i} + 3\mathbf{j} + \mathbf{k}$, $\mathbf{v} = -2\mathbf{i} - 3\mathbf{j} - 7\mathbf{k}$

10. $\mathbf{u} = -\mathbf{i} + 3\mathbf{k}$, $\mathbf{v} = 4\mathbf{j}$

11. $\mathbf{u} = -\mathbf{i} + 2\mathbf{j} - 3\mathbf{k}$, $\mathbf{v} = -\mathbf{i} - 3\mathbf{j} + 4\mathbf{k}$

12. Normalizar los vectores en los ejercicios del 6 al 8.

13. Hallar el ángulo entre los vectores en los ejercicios del 9 al 11. De ser necesario, expresar la respuesta en términos de $\cos^{-1}$.

14. Hallar la proyección de $\mathbf{u} = -\mathbf{i} + \mathbf{j} + \mathbf{k}$ sobre $\mathbf{v} = 2\mathbf{i} + \mathbf{j} - 3\mathbf{k}$.

15. Hallar la proyección de $\mathbf{v} = 2\mathbf{i} + \mathbf{j} - 3\mathbf{k}$ sobre $\mathbf{u} = -\mathbf{i} + \mathbf{j} + \mathbf{k}$.

16. ¿Qué restricciones se deben tener sobre b para que el vector $2\mathbf{i} + b\mathbf{j}$ sea ortogonal a (a) $-3\mathbf{i} + 2\mathbf{j} + \mathbf{k}$ y (b) $\mathbf{k}$?

17. Hallar dos vectores no paralelos, ambos ortogonales a $(1, 1, 1)$.

18. Hallar la recta que pasa por $(3, 1, -2)$ que interseca y es perpendicular a la recta $x = -1 + t$, $y = -2 + t$, $z = -1 + t$. [IDEA: Si (x_0, y_0, z_0) es el punto de intersección, hallar sus coordenadas.]

19. Suponer que una fuerza $\mathbf{F}$ (por ejemplo, la gravedad) actúa verticalmente hacia abajo sobre un objeto situado en un plano inclinado en un ángulo de 45° respecto a la horizontal. Expresar esta fuerza como suma de una fuerza que actúe paralela al plano y una que actúe perpendicular a él.

20. Suponer que un objeto moviéndose en la dirección de $\mathbf{i}+\mathbf{j}$ está bajo la acción de una fuerza dada por el vector $2\mathbf{i} + \mathbf{j}$. Expresar esta fuerza como una suma de una fuerza en la dirección del movimiento y una fuerza perpendicular a la dirección del movimiento.

21. Una fuerza de 6 N (newtons) forma un ángulo de $\pi/4$ radianes con el eje y, apuntando a la derecha. La fuerza actúa en contra del movimiento de un objeto a lo largo de la recta que une $(1, 2)$ con $(5, 4)$.

(a) Hallar una fórmula para el vector de fuerza $\mathbf{F}$.

(b) Hallar el ángulo θ entre la dirección del desplazamiento $\mathbf{D} = (5-1)\mathbf{i}+(4-2)\mathbf{j}$ y la dirección de la fuerza $\mathbf{F}$.

(c) El *trabajo* realizado es $\mathbf{F}\cdot\mathbf{D}$, o de manera equivalente, $\|\mathbf{F}\|\,\|\mathbf{D}\|\cos\theta$. Calcular el trabajo con ambas fórmulas y comparar los resultados.

***22.** Un fluido fluye a través de una superficie plana con vector de velocidad uniforme $\mathbf{v}$. Sea $\mathbf{n}$ una normal unitaria a la superficie del plano. Mostrar que $\mathbf{v}\cdot\mathbf{n}$ es el volumen del fluido que pasa por una unidad de área del plano en una unidad de tiempo.

1.3 EL PRODUCTO CRUZ

En la sección 1.2 hemos definido un producto de vectores que daba como resultado un escalar. En esta sección definiremos un producto de vectores que da como resultado un vector; esto es, mostraremos cómo dados dos vectores **a** y **b**, podemos producir un tercer vector $\mathbf{a} \times \mathbf{b}$, llamado el *producto cruz* de **a** y **b**. Este nuevo vector tendrá la muy agradable propiedad geométrica de ser perpendicular al plano generado (determinado) por **a** y **b**. La definición del producto cruz está basada en los conceptos de matriz y determinante que desarrollamos primero. Una vez hecho esto podremos estudiar las implicaciones geométricas de la estructura matemática construida.

Definimos una *matriz* de 2×2 como un arreglo

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

donde a_{11}, a_{12}, a_{21} y a_{22} son cuatro escalares. Por ejemplo,

$$\begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix}, \quad \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 13 & 7 \\ 6 & 11 \end{bmatrix}$$

son matrices de 2×2. El *determinante*

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

de dicha matriz es el número real definido por la ecuación

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{1}$$

EJEMPLO 1

$$\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} = 1 - 1 = 0; \quad \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 4 - 6 = -2; \quad \begin{vmatrix} 5 & 6 \\ 7 & 8 \end{vmatrix} = 40 - 42 = -2. \quad \blacktriangle$$

Una matriz de 3×3 es un arreglo

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

donde, de nuevo, cada a_{ij} es un escalar; a_{ij} denota el registro o posición en el arreglo que está en el i-ésimo renglón y la j-ésima columna. Definimos el

determinante de una matriz de 3 × 3 por la regla

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \quad (2)$$

Sería difícil memorizar la fórmula (2) sin algún recurso mnemotécnico. La regla que hay que aprender es que nos movemos a lo largo del primer renglón, multiplicando a_{1j} por el determinante de la matriz de 2×2 obtenida al eliminar el primer renglón y la j-ésima columna, y después sumando todo esto, pero recordando poner un signo de resta antes del término a_{12}. Por ejemplo, el determinante multiplicado por el término de enmedio en la fórmula (2), a saber

$$\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

se obtiene al eliminar el primer renglón y la segunda columna de la matriz dada de 3 × 3:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}.$$

EJEMPLO 2

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} - 0\begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} + 0\begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix} = 1$$

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix} = 1\begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 2\begin{vmatrix} 4 & 6 \\ 7 & 9 \end{vmatrix} + 3\begin{vmatrix} 4 & 5 \\ 7 & 8 \end{vmatrix} = -3 + 12 - 9 = 0. \quad \blacktriangle$$

Una importante propiedad de los determinantes es que al intercambiar dos renglones o dos columnas se cambia su signo. Para determinantes de 2 × 2, esto es una consecuencia de la definición. Para renglones tenemos

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

$$= -(a_{21}a_{12} - a_{11}a_{22}) = -\begin{vmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{vmatrix}$$

y para columnas,

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = -(a_{12}a_{21} - a_{11}a_{22}) = -\begin{vmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{vmatrix}.$$

Dejamos al lector verificar esta propiedad para el caso de 3×3. (Ver el ejercicio 1 al final de la sección.)

Una segunda propiedad fundamental de los determinantes es que podemos sacar como factor común a escalares de cualquier renglón o columna. Para determinantes de 2×2 esto significa

$$\begin{vmatrix} \alpha a_{11} & a_{12} \\ \alpha a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & \alpha a_{12} \\ a_{21} & \alpha a_{22} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} \alpha a_{11} & \alpha a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ \alpha a_{21} & \alpha a_{22} \end{vmatrix}.$$

De manera análoga, para determinantes de 3×3 tenemos

$$\begin{vmatrix} \alpha a_{11} & \alpha a_{12} & \alpha a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \alpha \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & \alpha a_{12} & a_{13} \\ a_{21} & \alpha a_{22} & a_{23} \\ a_{31} & \alpha a_{32} & a_{33} \end{vmatrix}$$

y así sucesivamente. Estos resultados se siguen de las definiciones. En particular, si cualquier renglón o columna está formado(a) por ceros, entonces el valor del determinante es cero.

Un tercer hecho fundamental acerca de los determinantes es el siguiente: si cambiamos un renglón (o columna) mediante la suma de otro renglón (o, respectivamente, columna), no cambia el valor del determinante. Para el caso de 2×2 esto significa que

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 + b_1 & a_2 + b_2 \\ b_1 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_2 \\ b_1 + a_1 & b_2 + a_2 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 + a_2 & a_2 \\ b_1 + b_2 & b_2 \end{vmatrix} = \begin{vmatrix} a_1 & a_1 + a_2 \\ b_1 & b_1 + b_2 \end{vmatrix}.$$

Para el caso de 3×3, esto significa que

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 + b_1 & a_2 + b_2 & a_3 + b_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 + a_2 & a_2 & a_3 \\ b_1 + b_2 & b_2 & b_3 \\ c_1 + c_2 & c_2 & c_3 \end{vmatrix}$$

y así sucesivamente. De nuevo, se puede probar esta propiedad usando la definición de determinante (ver el ejercicio 35).

EJEMPLO 3 *Suponer*

$$\mathbf{a} = \alpha\mathbf{b} + \beta\mathbf{c}; \text{ i.e., } \mathbf{a} = (a_1, a_2, a_3) = \alpha(b_1, b_2, b_3) + \beta(c_1, c_2, c_3)$$

Mostrar que

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = 0.$$

SOLUCIÓN Probaremos el caso $\alpha \neq 0$, $\beta \neq 0$. El caso $\alpha = 0 = \beta$ es trivial, y el caso en que exactamente uno de α, β es cero, es una modificación sencilla del caso que probamos. Usando las propiedades fundamentales de los determinantes, el determinante en cuestión es

$$\begin{vmatrix} \alpha b_1 + \beta c_1 & \alpha b_2 + \beta c_2 & \alpha b_3 + \beta c_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$$= -\frac{1}{\alpha} \begin{vmatrix} \alpha b_1 + \beta c_1 & \alpha b_2 + \beta c_2 & \alpha b_3 + \beta c_3 \\ -\alpha b_1 & -\alpha b_2 & -\alpha b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

(factorizando $-1/\alpha$ en el segundo renglón)

$$= \left(-\frac{1}{\alpha}\right)\left(-\frac{1}{\beta}\right) \begin{vmatrix} \alpha b_1 + \beta c_1 & \alpha b_2 + \beta c_2 & \alpha b_3 + \beta c_3 \\ -\alpha b_1 & -\alpha b_2 & -\alpha b_3 \\ -\beta c_1 & -\beta c_2 & -\beta c_3 \end{vmatrix}$$

(factorizando $-1/\beta$ en el tercer renglón)

$$= \frac{1}{\alpha\beta} \begin{vmatrix} \beta c_1 & \beta c_2 & \beta c_3 \\ -\alpha b_1 & -\alpha b_2 & -\alpha b_3 \\ -\beta c_1 & -\beta c_2 & -\beta c_3 \end{vmatrix}$$ (sumando el segundo renglón al primero)

$$= \frac{1}{\alpha\beta} \begin{vmatrix} 0 & 0 & 0 \\ -\alpha b_1 & -\alpha b_2 & -\alpha b_3 \\ -\beta c_1 & -\beta c_2 & -\beta c_3 \end{vmatrix}$$ (sumando el tercer renglón al primero)

$$= 0. \quad \blacktriangle$$

NOTA HISTÓRICA

Parece que en 1693 Leibniz inventó y usó los determinantes por primera vez, en relación con soluciones de ecuaciones lineales. Maclaurin y Cramer desarrollaron sus propiedades entre 1729 y 1750; en particular, mostraron que la solución del sistema de ecuaciones

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$
$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

es

$$x_1 = \frac{1}{\Delta} \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \qquad x_2 = \frac{1}{\Delta} \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix},$$

y

$$x_3 = \frac{1}{\Delta}\begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix}$$

donde

$$\Delta = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix},$$

hecho conocido como la *regla de Cramer*. Posteriormente, Vandermonde (1772) y Cauchy (1812), al tratar los determinantes como un tema aparte que merecía atención especial, desarrollaron el campo de manera más sistemática, con contribuciones de Laplace, Jacobi, y otros. A Lagrange (1775) se deben fórmulas para volúmenes de paralelepípedos en términos de determinantes. Las estudiaremos más adelante en esta sección. No obstante que durante el siglo diecinueve los matemáticos estudiaron matrices y determinantes, los temas se consideraban por separado. Para conocer toda la historia hasta 1900, ver T. Muir, *The Theory of Determinants in the Historical Order of Development*, reimpreso por Dover, New York, 1960.

Ahora que hemos enunciado las propiedades necesarias de los determinantes y estudiado su historia, estamos listos para proceder con el producto cruz de vectores. Sean $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ y $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ vectores en $\mathbf{R}^3$. El *producto cruz* de $\mathbf{a}$ y $\mathbf{b}$, denotado por $\mathbf{a} \times \mathbf{b}$, está definido como el vector

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}\mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\mathbf{k},$$

o, simbólicamente,

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$

Aunque sólo definimos los determinantes para arreglos de números *reales*, esta expresión formal que incluye *vectores* es una ayuda útil para recordar el producto cruz.

Notar que el producto cruz de dos vectores es otro vector; a veces se le llama *producto vectorial.*

EJEMPLO 4 *Hallar* $(3\mathbf{i} - \mathbf{j} + \mathbf{k}) \times (\mathbf{i} + 2\mathbf{j} - \mathbf{k})$.

SOLUCIÓN

$$(3\mathbf{i} - \mathbf{j} + \mathbf{k}) \times (\mathbf{i} + 2\mathbf{j} - \mathbf{k}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & -1 & 1 \\ 1 & 2 & -1 \end{vmatrix} = -\mathbf{i} + 4\mathbf{j} + 7\mathbf{k}. \quad \blacktriangle$$

Ciertas propiedades algebraicas del producto cruz se deducen de la definición. Si $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ son vectores y α, β y γ son escalares, entonces

(i) $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$

(ii) $\mathbf{a} \times (\beta\mathbf{b} + \gamma\mathbf{c}) = \beta(\mathbf{a} \times \mathbf{b}) + \gamma(\mathbf{a} \times \mathbf{c})$
$(\alpha\mathbf{a} + \beta\mathbf{b}) \times \mathbf{c} = \alpha(\mathbf{a} \times \mathbf{c}) + \beta(\mathbf{b} \times \mathbf{c})$

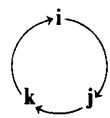
Notar que $\mathbf{a} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{a})$, por la propiedad (i). Así, $\mathbf{a} \times \mathbf{a} = \mathbf{0}$. En particular,

$$\mathbf{i} \times \mathbf{i} = 0, \qquad \mathbf{j} \times \mathbf{j} = 0, \qquad \mathbf{k} \times \mathbf{k} = 0.$$

Además

$$\mathbf{i} \times \mathbf{j} = \mathbf{k}, \qquad \mathbf{j} \times \mathbf{k} = \mathbf{i}, \qquad \mathbf{k} \times \mathbf{i} = \mathbf{j},$$

lo cual se puede recordar al permutar cíclicamente $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$ así:



Nuestro siguiente objetivo es proporcionar una interpretación geométrica del producto cruz. Para hacerlo, introducimos primero el triple producto. Dados tres vectores $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$, el número real

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$$

se llama el *triple producto* de $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ (en ese orden). Para obtener una fórmula sean $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ y $\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$. Entonces

$$\begin{aligned}\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) &= (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot \left(\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} \mathbf{k} \right) \\ &= a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.\end{aligned}$$

Esto se puede escribir de manera más concisa como

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Supongan ahora que $\mathbf{a}$ es un vector en el plano generado por los vectores $\mathbf{b}$ y $\mathbf{c}$. Esto significa que el primer renglón en la expresión como determinante de $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ es de la forma $\mathbf{a} = \alpha\mathbf{b} + \beta\mathbf{c}$, y por lo tanto $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = 0$, por el

ejemplo 3. En otras palabras, el vector $\mathbf{b} \times \mathbf{c}$ es ortogonal a cualquier vector en el plano generado por $\mathbf{b}$ y $\mathbf{c}$, en particular tanto a $\mathbf{b}$ como a $\mathbf{c}$.

A continuación calculamos la magnitud de $\mathbf{b} \times \mathbf{c}$. Noten que

$$\begin{aligned}\|\mathbf{b} \times \mathbf{c}\|^2 &= \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}^2 + \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix}^2 + \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}^2 \\ &= (b_2c_3 - b_3c_2)^2 + (b_1c_3 - c_1b_3)^2 + (b_1c_2 - c_1b_2)^2.\end{aligned}$$

Desarrollando esta última expresión, vemos que es igual a

$$\begin{aligned}&(b_1^2 + b_2^2 + b_3^2)(c_1^2 + c_2^2 + c_3^2) - (b_1c_1 + b_2c_2 + b_3c_3)^2 \\ &\quad = \|\mathbf{b}\|^2 \|\mathbf{c}\|^2 - (\mathbf{b} \cdot \mathbf{c})^2 = \|\mathbf{b}\|^2 \|\mathbf{c}\|^2 - \|\mathbf{b}\|^2 \|\mathbf{c}\|^2 \cos^2 \theta \\ &\quad = \|\mathbf{b}\|^2 \|\mathbf{c}\|^2 \operatorname{sen}^2 \theta\end{aligned}$$

donde θ es el ángulo entre $\mathbf{b}$ y $\mathbf{c}$, $0 \leq \theta \leq \pi$.

Combinando nuestros resultados concluimos que $\mathbf{b} \times \mathbf{c}$ *es un vector perpendicular al plano generado por* $\mathbf{b}$ *y* $\mathbf{c}$, *con longitud* $\|\mathbf{b}\| \|\mathbf{c}\| |\operatorname{sen} \theta|$. Sin embargo, hay dos vectores que pueden satisfacer estas condiciones, pues se pueden escoger dos direcciones que sean perpendiculares (o normales) al plano P generado por $\mathbf{b}$ y $\mathbf{c}$. Esto se ve claro en la figura 1.3.1, que muestra las dos posibilidades $\mathbf{n}_1$ y $-\mathbf{n}_1$ perpendiculares a P, con $\|\mathbf{n}_1\| = \|-\mathbf{n}_1\| = \|\mathbf{b}\| \|\mathbf{c}\| |\operatorname{sen} \theta|$.

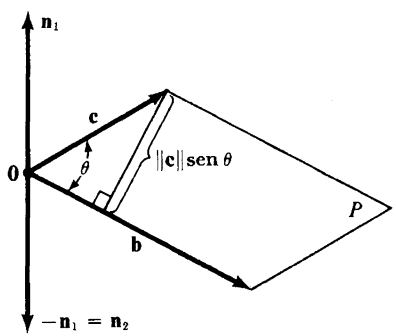


Figura 1.3.1 $\mathbf{n}_1$ y $\mathbf{n}_2$ son los dos posibles vectores ortogonales a $\mathbf{b}$ y a $\mathbf{c}$, ambos con norma $\|\mathbf{b}\| \|\mathbf{c}\| |\operatorname{sen} \theta|$.

¿Cuál es el vector que representa a $\mathbf{b} \times \mathbf{c}$?, ¿$\mathbf{n}_1$ o $-\mathbf{n}_1$? La respuesta es $\mathbf{n}_1 = \mathbf{b} \times \mathbf{c}$. Resuelvan algunos casos, como $\mathbf{k} = \mathbf{i} \times \mathbf{j}$, para verificarlo. La siguiente "regla de la mano derecha" determina la dirección de $\mathbf{b} \times \mathbf{c}$: Si colocan la palma de su mano derecha de manera que sus dedos se curven desde $\mathbf{b}$ en la dirección de $\mathbf{c}$ en un ángulo θ, el dedo pulgar apuntará en la dirección de $\mathbf{b} \times \mathbf{c}$ (figura 1.3.2).

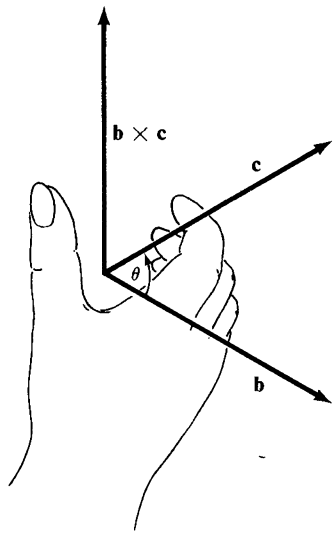


Figura 1.3.2 Regla de la mano derecha para determinar en cuál de las dos direcciones posibles apunta $\mathbf{b} \times \mathbf{c}$.

Si $\mathbf{b}$ y $\mathbf{c}$ son colineales, $\operatorname{sen}\theta = 0$, de modo que $\mathbf{b} \times \mathbf{c} = \mathbf{0}$. Si $\mathbf{b}$ y $\mathbf{c}$ no son colineales, entonces generan un plano y $\mathbf{b} \times \mathbf{c}$ es un vector perpendicular a este plano. *La longitud de* $\mathbf{b} \times \mathbf{c}$, $\|\mathbf{b}\| \|\mathbf{c}\| |\operatorname{sen}\theta|$, *es simplemente el área del paralelogramo que tiene como lados adyacentes a los vectores* $\mathbf{b}$ *y* $\mathbf{c}$ (figura 1.3.3).

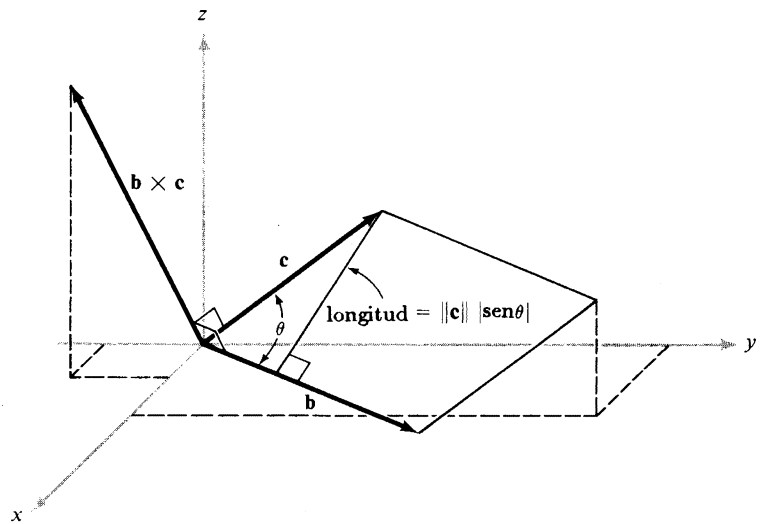


Figura 1.3.3 La longitud de $\mathbf{b} \times \mathbf{c}$ es igual al área del paralelogramo formado por $\mathbf{b}$ y $\mathbf{c}$.

EJEMPLO 5 *Hallar un vector unitario ortogonal a los vectores* $\mathbf{i}+\mathbf{j}$ *y* $\mathbf{j}+\mathbf{k}$.

SOLUCIÓN Un vector perpendicular a $\mathbf{i}+\mathbf{j}$ y a $\mathbf{j}+\mathbf{k}$ es el vector

$$(\mathbf{i}+\mathbf{j})\times(\mathbf{j}+\mathbf{k}) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{vmatrix} = \mathbf{i}-\mathbf{j}+\mathbf{k}.$$

Como $\|\mathbf{i}-\mathbf{j}+\mathbf{k}\| = \sqrt{3}$, el vector

$$\frac{1}{\sqrt{3}}(\mathbf{i}-\mathbf{j}+\mathbf{k})$$

es un vector unitario perpendicular a $\mathbf{i}+\mathbf{j}$ y $\mathbf{j}+\mathbf{k}$. ▲

Usando el producto cruz podemos obtener la interpretación geométrica básica de los determinantes de 2×2 y, más adelante, de 3×3. Sean $\mathbf{b} = b_1\mathbf{i}+b_2\mathbf{j}$ y $\mathbf{c} = c_1\mathbf{i}+c_2\mathbf{j}$ dos vectores en el plano. Si θ denota el ángulo entre $\mathbf{b}$ y $\mathbf{c}$, hemos visto que $\|\mathbf{b}\times\mathbf{c}\| = \|\mathbf{b}\|\,\|\mathbf{c}\|\,|\operatorname{sen}\theta|$. Como ya se dijo, $\|\mathbf{b}\|\,\|\mathbf{c}\|\,|\operatorname{sen}\theta|$ es el área del paralelogramo con lados adyacentes $\mathbf{b}$ y $\mathbf{c}$ (ver la figura 1.3.3). Usando la definición del producto cruz,

$$\mathbf{b}\times\mathbf{c} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ b_1 & b_2 & 0 \\ c_1 & c_2 & 0 \end{vmatrix} = \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}\mathbf{k}.$$

Así, $\|\mathbf{b}\times\mathbf{c}\|$ es el valor absoluto del determinante

$$\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} = b_1c_2 - b_2c_1.$$

De aquí se sigue que *el valor absoluto del determinante anterior es el área del paralelogramo que tiene como lados adyacentes a los vectores* $\mathbf{b} = b_1\mathbf{i}+b_2\mathbf{j}$ *y* $\mathbf{c} = c_1\mathbf{i}+c_2\mathbf{j}$.

EJEMPLO 6 *Hallar el área del triángulo con vértices en los puntos* $(1,1)$, $(0,2)$, *y* $(3,2)$ (figura 1.3.4).

SOLUCIÓN Sean $\mathbf{a} = \mathbf{i}+\mathbf{j}$, $\mathbf{b} = 2\mathbf{j}$ y $\mathbf{c} = 3\mathbf{i}+2\mathbf{j}$. Es claro que el triángulo cuyos vértices son los extremos de los vectores $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ tiene la misma área que el triángulo con vértices en $\mathbf{0}$, $\mathbf{b}-\mathbf{a}$ y $\mathbf{c}-\mathbf{a}$ (figura 1.3.4). En efecto, este último es sólo una traslación del triángulo anterior. Como el área de este triángulo trasladado es la mitad del área del paralelogramo con lados adyacentes $\mathbf{b}-\mathbf{a} = -\mathbf{i}+\mathbf{j}$ y $\mathbf{c}-\mathbf{a} = 2\mathbf{i}+\mathbf{j}$, hallamos que el área del triángulo con vértices $(1,1)$,

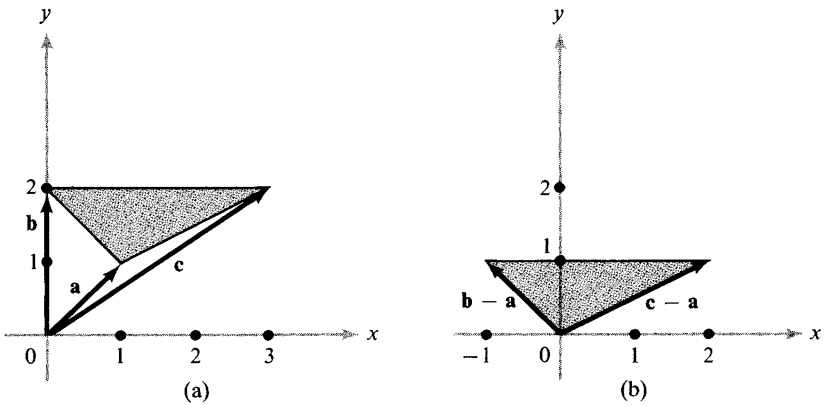


Figura 1.3.4 Problema (a): Hallar el área A del triángulo sombreado. Solución: Expresar los lados como diferencias de vectores (b) para obtener $A = \frac{1}{2}\|(\mathbf{b} - \mathbf{a}) \times (\mathbf{c} - \mathbf{a})\|$.

$(0, 2)$ y $(3, 2)$ es el valor absoluto de

$$\frac{1}{2}\begin{vmatrix} -1 & 1 \\ 2 & 1 \end{vmatrix} = -\frac{3}{2},$$

esto es, $\frac{3}{2}$. ▲

Hay una interpretación de los determinantes de matrices de 3×3 como volúmenes, que es análoga a la interpretación de los determinantes de matrices de 2×2 como áreas. Sean $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ y $\mathbf{c} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$, vectores en $\mathbf{R}^3$. Mostraremos que *el volumen del paralelepípedo con aristas adyacentes* **a**, **b** *y* **c** (figura 1.3.5) *es el valor absoluto del determinante*

$$D = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

Sabemos que $\|\mathbf{a} \times \mathbf{b}\|$ es el área del paralelogramo con lados adyacentes **a** y **b**. Más aún, $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}| = \|\mathbf{c}\|\,\|\mathbf{a} \times \mathbf{b}\| \cos\psi$, donde ψ es el ángulo agudo que forma **c** con la normal al plano generado por **a** y **b**. Como el volumen del paralelepípedo con aristas adyacentes **a**, **b** y **c** es el producto del área de la base $\|\mathbf{a} \times \mathbf{b}\|$ por la altura $\|\mathbf{c}\| \cos\psi$, se sigue que el volumen es $|(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}|$. Vimos en la pág. 35 que $D = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$. Al intercambiar renglones vemos que $D = -\mathbf{c} \cdot (\mathbf{b} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c}$; por lo tanto, el valor absoluto de D es el volumen del paralelepípedo con aristas adyacentes **a**, **b** y **c**.

Para concluir esta sección, usaremos métodos vectoriales para determinar la ecuación de un plano en el espacio. Sean P un plano en el espacio, **a** un vector que termina en el plano, y **n** un vector normal al plano (ver la figura 1.3.6).

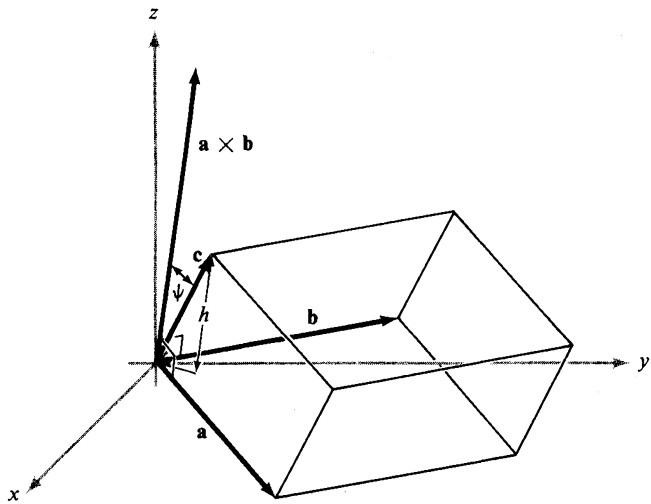


Figura 1.3.5 El volumen del paralelepípedo formado por **a**, **b**, **c** es el valor absoluto del determinante de la matriz de 3×3 con renglones **a**, **b** y **c**.

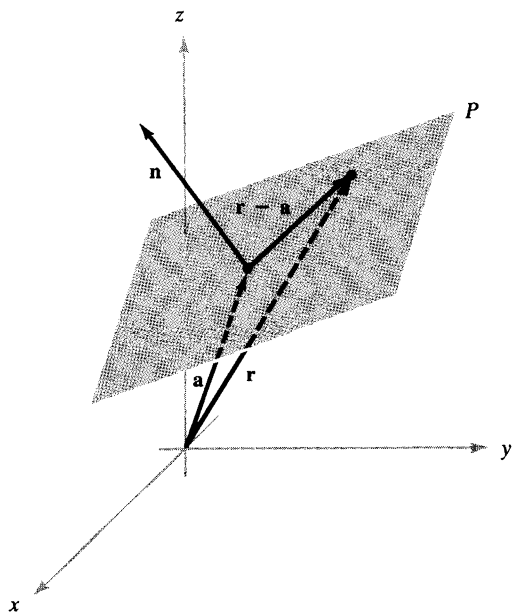


Figura 1.3.6 Los puntos **r** del plano que pasa por **a** y es perpendicular a **n** satisfacen la ecuación $(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0$.

Si $\mathbf{r}$ es un vector en $\mathbf{R}^3$, entonces el extremo de $\mathbf{r}$ está en el plano P si, y sólo si, $\mathbf{r}-\mathbf{a}$ es paralelo a P y, por lo tanto, si, y sólo si, $(\mathbf{r}-\mathbf{a})\cdot\mathbf{n} = 0$ ($\mathbf{n}$ es perpendicular a cualquier vector paralelo a P —ver la figura 1.3.6—). Como el producto interno es distributivo, esta última condición es equivalente a $\mathbf{r}\cdot\mathbf{n} = \mathbf{a}\cdot\mathbf{n}$. Por lo tanto, si hacemos $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{n} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ y $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, se sigue que el extremo de $\mathbf{r}$ está en P si, y sólo si,

$$Ax + By + Cz = \mathbf{r}\cdot\mathbf{n} = \mathbf{a}\cdot\mathbf{n} = Aa_1 + Ba_2 + Ca_3. \tag{3}$$

Como $\mathbf{n}$ y $\mathbf{a}$ se tomaron fijos, el lado derecho de la ecuación (3) es una constante, digamos, $-D$. Entonces una ecuación que determina el plano P es

$$Ax + By + Cz + D = 0. \tag{4}$$

donde $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ es normal a P; recíprocamente, si A, B y C no son cero simultáneamente, el conjunto de puntos (x, y, z) que satisface la ecuación (4) es un plano con normal $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$. La ecuación (4) es lineal en las tres variables x, y, z y así corresponde geométricamente a una superficie lineal, esto es, un plano, en $\mathbf{R}^3$.

Los cuatro números A, B, C, D no están determinados de manera única por P. Para verlo, noten que (x, y, z) satisface la ecuación (4) si, y sólo si, además satisface la relación

$$(\lambda A)x + (\lambda B)y + (\lambda C)z + (\lambda D) = 0$$

para cualquier constante $\lambda \neq 0$. Si A, B, C, D y A', B', C', D' determinan el mismo plano P, entonces $A = \lambda A'$, $B = \lambda B'$, $C = \lambda C'$, $D = \lambda D'$ para un escalar λ. Decimos que A, B, C, D *están determinadas por P salvo un múltiplo escalar.* Recíprocamente, dados A, B, C, D y A', B', C', D', determinan el mismo plano si $A = \lambda A'$, $B = \lambda B'$, $C = \lambda C'$, $D = \lambda D'$ para algún escalar λ. Este hecho se aclarará en el ejemplo 8.

El plano con normal $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$, que pasa por un punto $R = (x_0, y_0, z_0)$ es

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0 \tag{5}$$

(notar que $x = x_0$, $y = y_0$, $z = z_0$ satisface la ecuación (5), y entonces, en este caso, $D = -(Ax_0 + By_0 + Cz_0)$.

EJEMPLO 7 *Determinar la ecuación del plano perpendicular al vector* $\mathbf{i}+\mathbf{j}+\mathbf{k}$, *que contiene al punto* $(1, 0, 0)$.

SOLUCIÓN De la ecuación (5), el plano es $1(x - 1) + 1(y - 0) + 1(z - 0) = 0$; esto es, $x + y + z = 1$. ▲

EJEMPLO 8 *Hallar la ecuación del plano que contiene a los puntos* $(1,1,1)$, $(2,0,0)$ *y* $(1,1,0)$.

SOLUCIÓN *Método 1.* Cualquier ecuación del plano es de la forma $Ax + By + Cz + D = 0$. Como los puntos $(1,1,1)$ y $(2,0,0)$ y $(1,1,0)$ están en el plano, tenemos

$$\begin{aligned} A + B + C + D &= 0 \\ 2A \qquad\quad + D &= 0 \\ A + B \qquad + D &= 0 \end{aligned}$$

Mediante eliminación, reducimos este sistema de ecuaciones a la forma

$$\begin{aligned} 2A + D &= 0 \quad \text{(segunda ecuación)} \\ 2B + D &= 0 \quad (2 \times \text{tercera}-\text{segunda}) \\ C &= 0 \quad \text{(primera}-\text{tercera)} \end{aligned}$$

Como los números A, B C y D están determinados salvo un múltiplo escalar, podemos fijar el valor de uno y así los otros quedarán determinados de manera única. Si hacemos $D = -2$, entonces $A = +1$, $B = +1$, $C = 0$. Así, la ecuación del plano que contiene a los puntos dados es $x + y - 2 = 0$.

Método 2. Sean $\mathbf{a} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i}$ y $\mathbf{c} = \mathbf{i} + \mathbf{j}$. Cualquier vector normal al plano debe ser ortogonal a los vectores $\mathbf{a} - \mathbf{b}$ y $\mathbf{c} - \mathbf{b}$, que son paralelos al plano, ya que sus extremos están en el plano. Así, $\mathbf{n} = (\mathbf{a} - \mathbf{b}) \times (\mathbf{c} - \mathbf{b})$ es normal al plano. Al calcular el producto cruz tenemos,

$$\mathbf{n} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 1 & 1 \\ -1 & 1 & 0 \end{vmatrix} = -\mathbf{i} - \mathbf{j}.$$

Así, cualquier ecuación del plano es de la forma $-x - y + D = 0$ (salvo un múltiplo escalar). Como $(2,0,0)$ está en el plano, $D = +2$. Después de sustituir, obtenemos $x + y - 2 = 0$. ▲

EJEMPLO 9 *Determinar la distancia del punto* $\mathrm{E} = (x_1, y_1, z_1)$ *al plano con ecuación* $A(x - x_0) + B(y - y_0) + C(z - z_0) = Ax + By + Cz + D = 0$.

SOLUCIÓN Considerar al vector

$$\mathbf{n} = \frac{A\mathbf{i} + B\mathbf{j} + C\mathbf{k}}{\sqrt{A^2 + B^2 + C^2}},$$

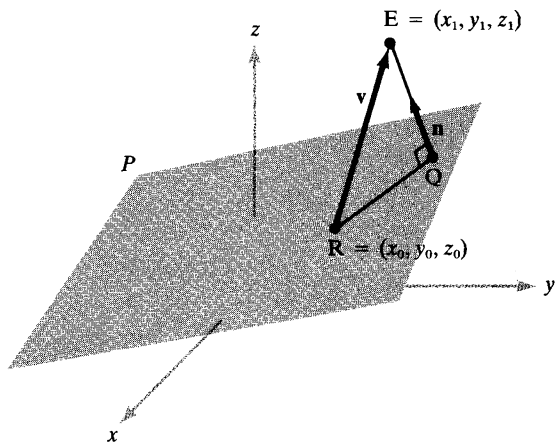


Figura 1.3.7 La geometría para determinar la distancia del punto E al plano P.

que es un vector unitario normal al plano. Bajar una perpendicular de E al plano y construir el triángulo REQ mostrado en la figura 1.3.7. La distancia $d = |\mathrm{EQ}|$ es la longitud de la proyección de $\mathbf{v} = \overrightarrow{\mathrm{RE}}$ (el vector de R a E) sobre $\mathbf{n}$; así,

$$\begin{aligned}\text{distancia} = |\mathbf{v} \cdot \mathbf{n}| &= |[(x_1 - x_0)\mathbf{i} + (y_1 - y_0)\mathbf{j} + (z_1 - z_0)\mathbf{k}] \cdot \mathbf{n}| \\ &= \frac{|A(x_1 - x_0) + B(y_1 - y_0) + C(z_1 - z_0)|}{\sqrt{A^2 + B^2 + C^2}}.\end{aligned}$$

Si el plano está dado en la forma $Ax + By + Cz + D = 0$, escogemos un punto (x_0, y_0, z_0) sobre él y notamos que $D = -(Ax_o + By_0 + Cz_0)$. Al sustituir en la fórmula anterior da

$$\text{distancia} = \frac{|Ax_1 + By_1 + Cz_1 + D|}{\sqrt{A^2 + B^2 + C^2}}. \quad \blacktriangle$$

EJERCICIOS

1. Verificar que al intercambiar dos renglones o dos columnas del determinante de 3×3

$$\begin{vmatrix} 1 & 2 & 1 \\ 3 & 0 & 1 \\ 2 & 0 & 2 \end{vmatrix}$$

se cambia el signo del determinante (escoger cualesquiera dos renglones o cualesquiera dos columnas).

2. Evaluar

(a) $\begin{vmatrix} 2 & -1 & 0 \\ 4 & 3 & 2 \\ 3 & 0 & 1 \end{vmatrix}$ (b) $\begin{vmatrix} 36 & 18 & 17 \\ 45 & 24 & 20 \\ 3 & 5 & -2 \end{vmatrix}$

(c) $\begin{vmatrix} 1 & 4 & 9 \\ 4 & 9 & 16 \\ 9 & 16 & 25 \end{vmatrix}$ (d) $\begin{vmatrix} 2 & 3 & 5 \\ 7 & 11 & 13 \\ 17 & 19 & 23 \end{vmatrix}$

3. Calcular $\mathbf{a} \times \mathbf{b}$, donde $\mathbf{a} = \mathbf{i} - 2\mathbf{j} + \mathbf{k}$, $\mathbf{b} = 2\mathbf{i} + \mathbf{j} + \mathbf{k}$.

4. Calcular $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$, donde $\mathbf{a}$ y $\mathbf{b}$ son como en el ejercicio 3 y $\mathbf{c} = 3\mathbf{i} - \mathbf{j} + 2\mathbf{k}$.

5. Hallar el área del paralelogramo que tiene como lados a los vectores $\mathbf{a}$ y $\mathbf{b}$ dados en el ejercicio 3.

6. Un triángulo tiene vértices $(0, 0, 0)$, $(1, 1, 1)$ y $(0, -2, 3)$. Hallar su área.

7. ¿Cuál es el volumen del paralelepípedo con aristas $2\mathbf{i} + \mathbf{j} - \mathbf{k}$, $5\mathbf{i} - 3\mathbf{k}$ e $\mathbf{i} - 2\mathbf{j} + \mathbf{k}$?

8. ¿Cuál es el volumen del paralelepípedo con aristas $\mathbf{i}$, $3\mathbf{j} - \mathbf{k}$ y $4\mathbf{i} + 2\mathbf{j} - \mathbf{k}$?

En los ejercicios del 9 al 12, describir todos los vectores unitarios ortogonales a los vectores dados.

9. $\mathbf{i}$, $\mathbf{j}$

10. $-5\mathbf{i} + 9\mathbf{j} - 4\mathbf{k}$, $7\mathbf{i} + 8\mathbf{j} + 9\mathbf{k}$

11. $-5\mathbf{i} - 9\mathbf{j} - 4\mathbf{k}$, $7\mathbf{i} + 8\mathbf{j} + 9\mathbf{k}$, $\mathbf{0}$

12. $2\mathbf{i} - 4\mathbf{j} + 3\mathbf{k}$, $-4\mathbf{i} + 8\mathbf{j} - 6\mathbf{k}$

13. Calcular $\mathbf{u} + \mathbf{v}$, $\mathbf{u} \cdot \mathbf{v}$, $\|\mathbf{u}\|$, $\|\mathbf{v}\|$ y $\mathbf{u} \times \mathbf{v}$ donde $\mathbf{u} = \mathbf{i} - 2\mathbf{j} + \mathbf{k}$, $\mathbf{v} = 2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$.

14. Repetir el ejercicio 13 para $\mathbf{u} = 3\mathbf{i} + \mathbf{j} - \mathbf{k}$, $\mathbf{v} = -6\mathbf{i} - 2\mathbf{j} - 2\mathbf{k}$.

15. Hallar una ecuación para el plano que

(a) es perpendicular a $\mathbf{v} = (1, 1, 1)$ y pasa por $(1, 0, 0)$.

(b) es perpendicular a $\mathbf{v} = (1, 2, 3)$ y pasa por $(1, 1, 1)$.

(c) es perpendicular a la recta $\mathbf{l}(t) = (5, 0, 2)t + (3, -1, 1)$ y pasa por $(5, -1, 0)$.

(d) es perpendicular a la recta $\mathbf{l}(t) = (-1, -2, 3)t + (0, 7, 1)$ y pasa por $(2, 4, -1)$.

16. Hallar una ecuación para el plano que pasa por (a) $(0, 0, 0)$, $(2, 0, -1)$ y $(0, 4, -3)$; (b) $(1, 2, 0)$, $(0, 1, -2)$ y $(4, 0, 1)$; y (c) $(2, -1, 3)$, $(0, 0, 5)$ y $(5, 7, -1)$.

17. (a) Probar las identidades del triple producto vectorial $(\mathbf{A} \times \mathbf{B}) \times \mathbf{C} = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{B} \cdot \mathbf{C})\mathbf{A}$ y $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = (\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}$.

(b) Probar $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times (\mathbf{v} \times \mathbf{w})$ si, y sólo si $(\mathbf{u} \times \mathbf{w}) \times \mathbf{v} = \mathbf{0}$.

(c) Probar $(\mathbf{u} \times \mathbf{v}) \times \mathbf{w} + (\mathbf{v} \times \mathbf{w}) \times \mathbf{u} + (\mathbf{w} \times \mathbf{u}) \times \mathbf{v} = \mathbf{0}$ (la *identidad de Jacobi*).

18. (a) Probar sin recursos geométricos, que

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \mathbf{v} \cdot (\mathbf{w} \times \mathbf{u}) = \mathbf{w} \cdot (\mathbf{u} \times \mathbf{v}) = -\mathbf{u} \cdot (\mathbf{w} \times \mathbf{v})$$
$$= -\mathbf{w} \cdot (\mathbf{v} \times \mathbf{u}) = -\mathbf{v} \cdot (\mathbf{u} \times \mathbf{w})$$

(b) Probar

$$(\mathbf{u} \times \mathbf{v}) \cdot (\mathbf{u}' \times \mathbf{v}') = (\mathbf{u} \cdot \mathbf{u}')(\mathbf{v} \cdot \mathbf{v}') - (\mathbf{u} \cdot \mathbf{v}')(\mathbf{u}' \cdot \mathbf{v}) = \begin{vmatrix} \mathbf{u} \cdot \mathbf{u}' & \mathbf{u} \cdot \mathbf{v}' \\ \mathbf{u}' \cdot \mathbf{v} & \mathbf{v} \cdot \mathbf{v}' \end{vmatrix}$$

[IDEA: Usar la parte (a) y el ejercicio 17(a).]

19. Verificar la regla de Cramer presentada en la nota histórica de la página 33.

20. Hallar una ecuación para el plano que pasa por $(2, -1, 3)$ y es perpendicular a $\mathbf{v} = (1, -2, 2) + t(3, -2, 4)$.

21. Hallar una ecuación para el plano que pasa por $(1, 2, -3)$ y es perpendicular a $\mathbf{v} = (0, -2, 1) + t(1, -2, 3)$.

22. Hallar la ecuación de la recta que pasa por $(1, -2, -3)$ y es perpendicular al plano $3x - y - 2z + 4 = 0$.

23. Hallar una ecuación para el plano que contiene a las dos rectas

$$\mathbf{v}_1 = (0, 1, -2) + t(2, 3, -1) \qquad \text{y} \qquad \mathbf{v}_2 = (2, -1, 0) + t(2, 3, -1).$$

24. Hallar la distancia de $(2, 1, -1)$ al plano $x - 2y + 2z + 5 = 0$.

25. Hallar una ecuación del plano que contiene a la recta $\mathbf{v} = (-1, 1, 2) + t(3, 2, 4)$ y es perpendicular al plano $2x + y - 3z + 4 = 0$.

26. Hallar una ecuación del plano que pasa por $(3, 2, -1)$ y $(1, -1, 2)$ y es paralelo a la recta $\mathbf{v} = (1, -1, 0) + t(3, 2, -2)$.

27. Rehacer los ejercicios 19 y 20 de la sección 1.1 usando el producto punto y los conocimientos adquiridos acerca de normales y planos.

28. Dados los vectores $\mathbf{a}$ y $\mathbf{b}$, ¿es cierto que las ecuaciones $\mathbf{x} \times \mathbf{a} = \mathbf{b}$ y $\mathbf{x} \cdot \mathbf{a} = \|\mathbf{a}\|$ determinan sólo un vector $\mathbf{x}$? Dar argumentos geométricos y analíticos.

29. Determinar la distancia del plano $12x + 13y + 5z + 2 = 0$ al punto $(1, 1, -5)$.

30. Hallar la distancia al punto $(6, 1, 0)$ del plano que pasa por el origen y es perpendicular a $\mathbf{i} - 2\mathbf{j} + \mathbf{k}$.

31. En mecánica se define el *momento* M *de una fuerza* $\mathbf{F}$ *alrededor de un punto* O como la magnitud de $\mathbf{F}$ por la distancia perpendicular d de O a la linea de acción de $\mathbf{F}$. (Recordemos del ejemplo 10, sección 1.1, que las fuerzas se pueden considerar vectores.) El *momento vector* $\mathbf{M}$ es el vector de magnitud M cuya dirección es perpendicular al plano de O y $\mathbf{F}$, determinado por la regla de la mano derecha. Mostrar que $\mathbf{M} = \mathbf{R} \times \mathbf{F}$, donde $\mathbf{R}$ es cualquier vector que va de O a la línea de acción de $\mathbf{F}$ (ver la figura 1.3.8.).

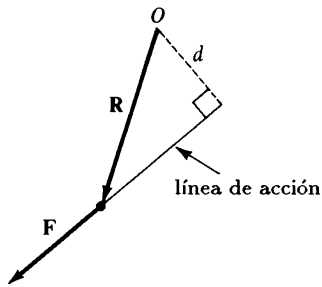


Figura 1.3.8 Momento de una fuerza.

32. La *velocidad angular* de rotación ω, de un cuerpo rígido tiene la misma dirección que el eje de rotación y magnitud igual a la tasa de giro en radianes por segundo. El sentido de ω se determina por la regla de la mano derecha.

(a) Sea $\mathbf{r}$ un vector que va del eje a un punto P del cuerpo rígido. Mostrar que el vector $\mathbf{v} = \omega \times \mathbf{r}$ es la velocidad de P, como en la figura 1.3.9, con $\omega = \mathbf{v}_1$ y $\mathbf{r} = \mathbf{v}_2$.

(b) Interpretar el resultado para la rotación de un carrusel alrededor de su eje, donde P es un punto en la circunferencia.

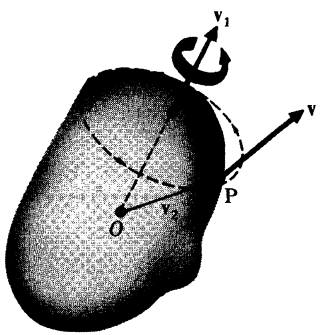


Figura 1.3.9 El punto P tiene vector velocidad $\mathbf{v}$.

33. Dos medios físicos con índices de refracción n_1 y n_2 están separados por una superficie plana perpendicular al vector unitario $\mathbf{N}$. Sean $\mathbf{a}$ y $\mathbf{b}$ vectores unitarios a lo largo de los rayos incidente y refractado, respectivamente, sus direcciones son las de dichos rayos de luz. Mostrar que $n_1(\mathbf{N} \times \mathbf{a}) = n_2(\mathbf{N} \times \mathbf{b})$, usando la *ley de Snell*, $\operatorname{sen}\theta_1 / \operatorname{sen}\theta_2 = n_2/n_1$, donde θ_1 y θ_2 son los ángulos de incidencia y refracción, repectivamente (ver la figura 1.3.10.)

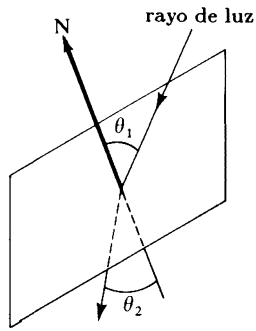


Figura 1.3.10 Ley de Snell.

*34. Justificar los pasos en los siguientes cálculos:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 7 & 8 & 10 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -11 \end{vmatrix} = \begin{vmatrix} -3 & -6 \\ -6 & -11 \end{vmatrix} = 33 - 36 = -3.$$

*35. Mostrar que, en una matriz, al sumar un múltiplo del primer renglón el segundo no se altera el determinante, esto es,

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 + \lambda a_1 & b_2 + \lambda b_1 & c_2 + \lambda c_1 \\ a_3 & b_3 & c_3 \end{vmatrix} = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}.$$

[De hecho, al sumar en una matriz un múltiplo de cualquier renglón (columna) a otro renglón (columna), no se altera el determinante.]

1.4 COORDENADAS ESFÉRICAS Y CILÍNDRICAS

La manera usual de representar un punto en el plano $\mathbf{R}^2$ es mediante las coordenadas rectangulares (x, y). Sin embargo, como ya seguramente lo aprendió el lector en cálculo elemental, las coordenadas polares en el plano pueden ser muy útiles. Como se muestra en la figura 1.4.1, las coordenadas (r, θ) están relacionadas con (x, y) mediante las fórmulas

$$x = r\cos\theta \qquad \text{y} \qquad y = r\operatorname{sen}\theta,$$

donde usualmente tomamos $r \geq 0$ y $0 \leq \theta < 2\pi$.

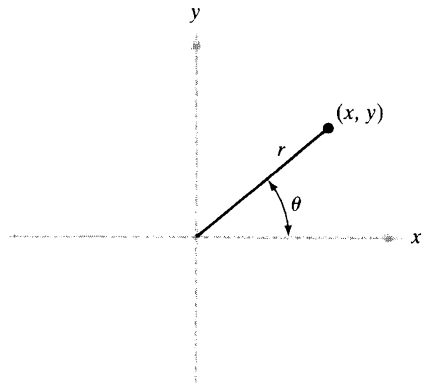


Figura 1.4.1 Las coordenadas polares de (x, y) son (r, θ).

A los lectores no familiarizados con las coordenadas polares se les recomienda estudiar las secciones respectivas en su libro de cálculo. Ahora vamos a exponer dos maneras de representar puntos en el espacio, además de las coordenadas cartesianas rectangulares (x, y, z). Estos sistemas coordenados alternativos son particularmente adecuados para ciertos tipos de problemas, como por ejemplo, la evaluación de integrales (ver la sección 6.3).

DEFINICIÓN *(ver la figura 1.4.2). Las* ***coordenadas cilíndricas*** (r, θ, z) *de un punto* (x, y, z) *están definidas por*

$$x = r\cos\theta, \qquad y = r\,\text{sen}\,\theta, \qquad z = z \tag{1}$$

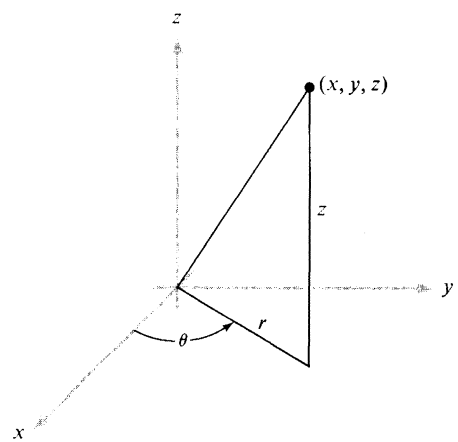


Figura 1.4.2 Representación de un punto (x, y, z) en términos de sus coordenadas cilíndricas r, θ y z.

o, explícitamente,

$$r = \sqrt{x^2 + y^2}, \qquad z = z, \qquad \theta = \begin{cases} \tan^{-1}(y/x) \text{ si } x > 0 \text{ y } y \geq 0 \\ \pi + \tan^{-1}(y/x) \text{ si } x < 0 \\ 2\pi + \tan^{-1}(y/x) \text{ si } x > 0 \text{ y } y < 0 \end{cases}$$

donde $\tan^{-1}(y/x)$ *está entre* $-\pi/2$ *y* $\pi/2$. *Si* $x = 0$, *entonces* $\theta = \pi/2$ *para* $y > 0$ *y* $3\pi/2$ *para* $y < 0$. *Si* $x = y = 0$, θ *no está definido.*

En otras palabras, para cada punto (x, y, z) representamos la primera y segunda coordenadas en términos de coordenadas polares y no alteramos la tercera. La fórmula (1) muestra que, dados (r, θ, z), la terna (x, y, z) está completamente determinada y, viceversa, si restringimos θ al intervalo $[0, 2\pi)$ (a veces es conveniente la extensión $(-\pi, \pi]$) y requerimos que $r > 0$.

Para ver por qué usamos el término "coordenadas cilíndricas", nótese que si $0 \leq \theta < 2\pi$, $-\infty < z < \infty$ y $r = a$ es una constante positiva, entonces el lugar geométrico de estos puntos es un cilindro de radio a (ver la figura 1.4.3).

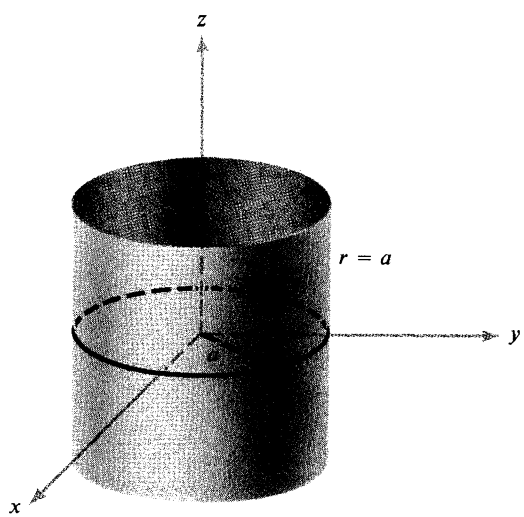


Figura 1.4.3 La gráfica de los puntos cuyas coordenadas cilíndricas satisfacen $r = a$ es un cilindro.

EJEMPLO 1 (a) *Hallar las coordenadas cilíndricas de* $(6, 6, 8)$ *y localizar el punto.* (b) *Si un punto tiene coordenadas cilíndricas* $(8, 2\pi/3, -3)$, *¿cuáles son sus coordenadas cartesianas? Localizarlo.*

SOLUCIÓN Para la parte (a), tenemos $r = \sqrt{6^2 + 6^2} = 6\sqrt{2}$ y $\theta = \tan^{-1}(6/6) = \tan^{-1}(1) = \pi/4$. Así, las coordenadas cilíndricas son $(6\sqrt{2}, \pi/4, 8)$. Éste es el

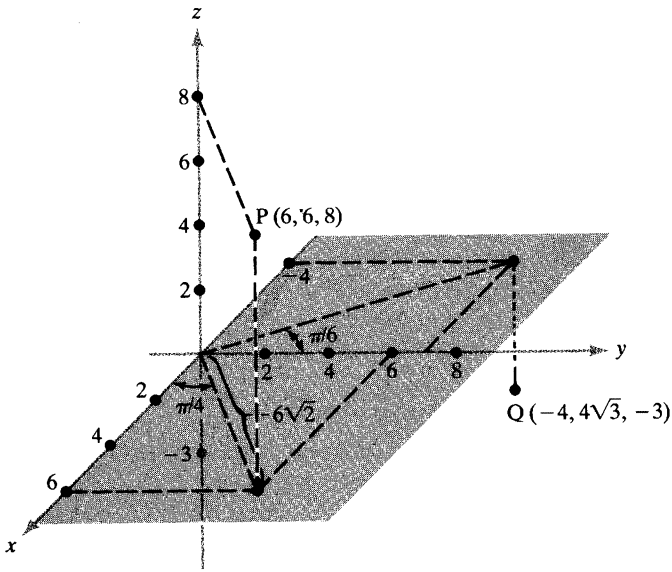


Figura 1.4.4 Ejemplos de conversión entre coordenadas cartesianas y cilíndricas.

punto P de la figura 1.4.4. Para la parte (b), tenemos

$$x = r\cos\theta = 8\cos\frac{2\pi}{3} = -\frac{8}{2} = -4$$

y

$$y = r\,\text{sen}\,\theta = 8\,\text{sen}\,\frac{2\pi}{3} = 8\frac{\sqrt{3}}{2} = 4\sqrt{3}.$$

Así, las coordenadas cartesianas son $(-4, 4\sqrt{3}, -3)$. Éste es el punto Q de la figura. ▲

Las coordenadas cilíndricas no son las únicas generalizaciones posibles de las coordenadas polares a tres dimensiones. Recuerden que en dos dimensiones la magnitud del vector $x\mathbf{i}+y\mathbf{j}$ (esto es, $\sqrt{x^2+y^2}$) es la r en el sistema de coordenadas polares. Para las coordenadas cilíndricas, la longitud del vector $x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$, a saber,

$$\rho = \sqrt{x^2+y^2+z^2},$$

no es una de las coordenadas del sistema (usamos sólo la magnitud $r = \sqrt{x^2+y^2}$, el ángulo θ y la "altura" z.)

Ahora modificaremos esto introduciendo el sistema de *coordenadas esféricas*, que usa a ρ como coordenada. Las coordenadas esféricas suelen ser útiles para resolver problemas donde hay simetría esférica (simetría alrededor de un punto), mientras que las coordenadas cilíndricas se pueden aplicar donde haya simetría cilíndrica (simetría alrededor de una recta).

Dado un punto $(x, y, z) \in \mathbf{R}^3$, sea

$$\rho = \sqrt{x^2 + y^2 + z^2}$$

y representemos x y y mediante coordenadas polares en el plano xy:

$$x = r\cos\theta, \qquad y = r\,\mathrm{sen}\,\theta \tag{2}$$

donde $r = \sqrt{x^2 + y^2}$ y θ está dada por la fórmula (1). La coordenada z está dada por

$$z = \rho\cos\phi,$$

donde ϕ es el ángulo (entre 0 y π, inclusive) que forma el radio vector $\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ con el eje z, en el plano que contiene al vector $\mathbf{v}$ y al eje z (ver la figura 1.4.5). Usando el producto punto podemos expresar ϕ como sigue:

$$\cos\phi = \frac{\mathbf{v}\cdot\mathbf{k}}{\|\mathbf{v}\|}, \qquad \text{i.e.,} \qquad \phi = \cos^{-1}\left(\frac{\mathbf{v}\cdot\mathbf{k}}{\|\mathbf{v}\|}\right).$$

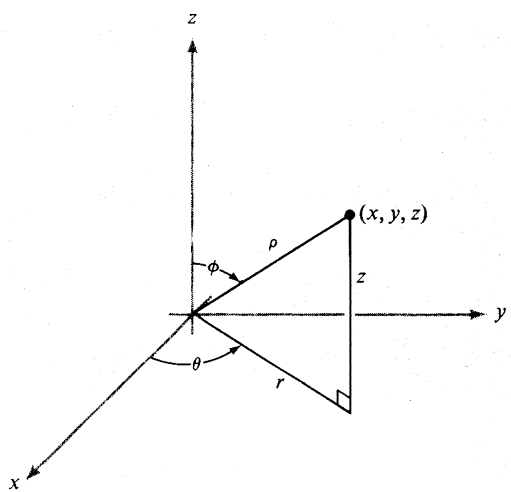


Figura 1.4.5 Coordenadas esféricas (ρ, θ, ϕ); la gráfica de los puntos que satisfacen $\rho = a$ es una esfera.

Tomamos como coordenadas las cantidades ρ, θ, ϕ. Como

$$r = \rho \operatorname{sen} \phi$$

podemos usar la fórmula (2) para expresar x, y y z en términos de coordenadas esféricas ρ, θ, ϕ.

DEFINICIÓN *Las* ***coordenadas esféricas*** *de* (x, y, z) *se definen como sigue:*

$$x = \rho \operatorname{sen} \phi \cos \theta, \qquad y = \rho \operatorname{sen} \phi \operatorname{sen} \theta, \qquad z = \rho \cos \phi \tag{3}$$

donde

$$\rho \geq 0, \qquad 0 \leq \theta < 2\pi, \qquad 0 \leq \phi \leq \pi.$$

Nótese que las coordenadas esféricas θ y ϕ se parecen a las coordenadas geográficas de longitud y latitud si consideramos al eje de la Tierra como el eje z. Sin embargo, hay diferencias: la longitud geográfica es $|\theta|$ y se llama longitud este u oeste, dependiendo de si θ es positivo o negativo; la latitud geográfica es $|\pi/2 - \phi|$ y se llama latitud norte o sur, dependiendo de si $\pi/2 - \theta$ es positivo o negativo.

Nótese que en coordenadas esféricas la ecuación de la esfera de radio a con centro en el origen toma la forma particularmente sencilla

$$\rho = a.$$

EJEMPLO 2

(a) *Hallar las coordenadas esféricas de* $(1, -1, 1)$ *y localizarlo.*

(b) *Hallar las coordenadas cartesianas de* $(3, \pi/6, \pi/4)$ *y localizarlo.*

(c) *Sea un punto con coordenadas cartesianas* $(2, -3, 6)$. *Hallar sus coordenadas esféricas y localizarlo.*

(d) *Sea un punto con coordenadas esféricas* $(1, -\pi/2, \pi/4)$. *Hallar sus coordenadas cartesianas y localizarlo.*

SOLUCIÓN

(a) $\rho = \sqrt{x^2 + y^2 + z^2} = \sqrt{1^2 + (-1)^2 + 1^2} = \sqrt{3},$

$$\theta = \tan^{-1}\left(\frac{y}{x}\right) = \tan^{-1}\left(\frac{-1}{1}\right) = -\frac{\pi}{4},$$

$$\phi = \cos^{-1}\left(\frac{z}{\rho}\right) = \cos^{-1}\left(\frac{1}{\sqrt{3}}\right) \approx 0.955 \approx 54.74°.$$

Ver la figura 1.4.6(a).

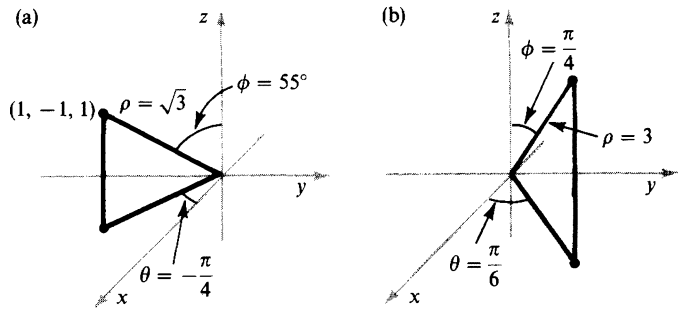


Figura 1.4.6 Búsqueda de (a) las coordenadas esféricas del punto $(1, -1, 1)$ y (b) las coordenadas cartesianas de $(3, \pi/6, \pi/4)$.

(b) $x = \rho \operatorname{sen} \phi \cos \theta = 3 \operatorname{sen} \left(\frac{\pi}{4}\right) \cos \left(\frac{\pi}{6}\right) = 3 \left(\frac{1}{\sqrt{2}}\right) \frac{\sqrt{3}}{2} = \frac{3\sqrt{3}}{2\sqrt{2}},$

$y = \rho \operatorname{sen} \phi \operatorname{sen} \theta = 3 \operatorname{sen} \left(\frac{\pi}{4}\right) \operatorname{sen} \left(\frac{\pi}{6}\right) = 3 \left(\frac{1}{\sqrt{2}}\right) \left(\frac{1}{2}\right) = \frac{3}{2\sqrt{2}},$

$z = \rho \cos \phi = 3 \cos \left(\frac{\pi}{4}\right) = \frac{3}{\sqrt{2}} = \frac{3\sqrt{2}}{2}.$

Ver la figura 1.4.6(b).

(c) $\rho = \sqrt{x^2 + y^2 + z^2} = \sqrt{2^2 + (-3)^2 + 6^2} = \sqrt{49} = 7,$

$\theta = \tan^{-1} \left(\frac{y}{x}\right) = \tan^{-1} \left(\frac{-3}{2}\right) \approx -0.983 \approx -56.31°,$

$\phi = \cos^{-1} \left(\frac{z}{\rho}\right) = \cos^{-1} \left(\frac{6}{7}\right) \approx 0.541 \approx 31.0°.$

Ver la figura 1.4.7(a).

(d) $x = \rho \operatorname{sen} \phi \cos \theta = 1 \operatorname{sen} \left(\frac{\pi}{4}\right) \cos \left(\frac{-\pi}{2}\right) = \left(\frac{\sqrt{2}}{2}\right) \cdot 0 = 0,$

$y = \rho \operatorname{sen} \phi \operatorname{sen} \theta = 1 \operatorname{sen} \left(\frac{\pi}{4}\right) \operatorname{sen} \left(\frac{-\pi}{2}\right) = \left(\frac{\sqrt{2}}{2}\right) (-1) = -\frac{\sqrt{2}}{2},$

$z = \rho \cos \phi = 1 \cos \left(\frac{\pi}{4}\right) = \frac{\sqrt{2}}{2}.$

Ver la figura 1.4.7(b). ▲

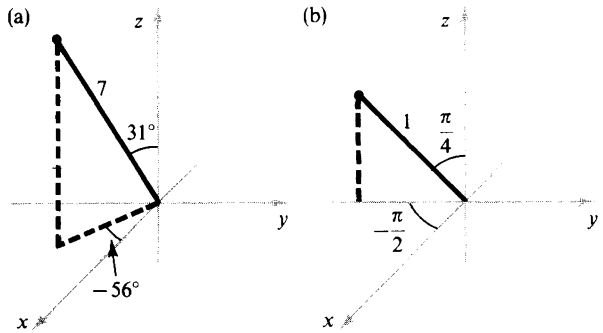


Figura 1.4.7 Búsqueda de (a) las coordenadas esféricas de $(2, -3, 6)$ y (b) las coordenadas cartesianas de $(1, -\pi/2, \pi/4)$.

EJEMPLO 3 *Expresar* (a) *la superficie* $xz = 1$ *y* (b) *la superficie* $x^2 + y^2 - z^2 = 1$ *en coordenadas esféricas.*

SOLUCIÓN De la fórmula (3), $x = \rho \operatorname{sen} \phi \cos \theta$, $z = \rho \cos \phi$, y, por lo tanto, la superficie (a) está formada por todos los (ρ, θ, ϕ) tales que

$$\rho^2 \operatorname{sen} \phi \cos \theta \cos \phi = 1, \qquad \text{i.e.,} \qquad \rho^2 \operatorname{sen} 2\phi \cos \theta = 2.$$

Para la parte (b) podemos escribir

$$x^2 + y^2 - z^2 = x^2 + y^2 + z^2 - 2z^2 = \rho^2 - 2\rho^2 \cos^2 \phi,$$

de manera que la superficie es $\rho^2(1 - 2\cos^2 \phi) = 1$, o bien $-\rho^2 \cos(2\phi) = 1$. ▲

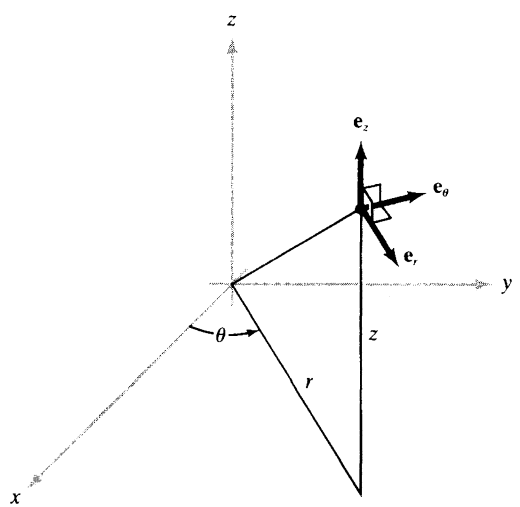


Figura 1.4.8 Vectores ortonormales $\mathbf{e}_r$, $\mathbf{e}_\theta$ y $\mathbf{e}_z$ asociados con las coordenadas cilíndricas. El vector $\mathbf{e}_r$ es paralelo a la recta denominada r.

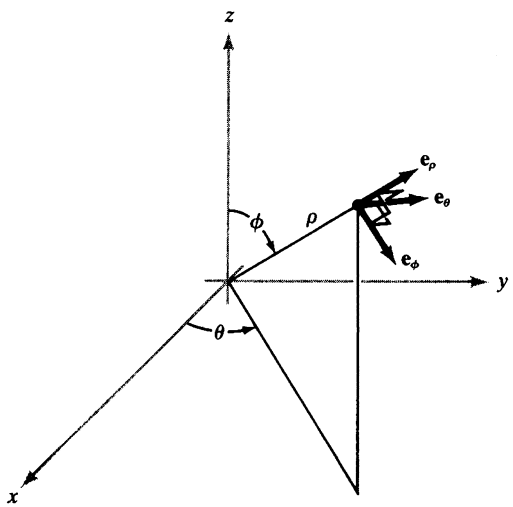


Figura 1.4.9 Vectores ortonormales $\mathbf{e}_\rho$, $\mathbf{e}_\theta$ y $\mathbf{e}_\phi$ asociados con las coordenadas esféricas.

Hay vectores unitarios asociados con las coordenadas cilíndricas y esféricas, que son la contraparte de $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$ para las coordenadas rectangulares. Se muestran en las figuras 1.4.8 y 1.4.9. Por ejemplo, $\mathbf{e}_r$ es el vector unitario paralelo al plano xy, en dirección radial, de manera que $\mathbf{e}_r = (\cos\theta)\mathbf{i} + (\text{sen}\,\theta)\mathbf{j}$. De manera análoga, en coordenadas esféricas $\mathbf{e}_\phi$ es el vector unitario tangente a la curva parametrizada por la variable ϕ, manteniendo fijas las variables ρ y θ. Usaremos estos vectores unitarios más adelante, cuando en cálculos vectoriales se utilicen coordenadas cilíndricas y esféricas (ver la sección 3.5).

EJERCICIOS

1. (a) Los puntos siguientes están dados en coordenadas cilíndricas; expresar cada uno en coordenadas rectangulares y coordenadas esféricas: $(1, 45°, 1)$, $(2, \pi/2, -4)$, $(0, 45°, 10)$, $(3, \pi/6, 4)$, $(1, -\pi/6, 0)$, $(2, 3\pi/4, -2)$. (Sólo el primer punto se resuelve en la Guía de estudio.)

(b) Cambiar los puntos siguientes de coordenadas rectangulares a coordenadas esféricas y a coordenadas cilíndricas: $(2, 1, -2)$, $(0, 3, 4)$, $(\sqrt{2}, 1, 1)$, $(-2\sqrt{3}, -2, 3)$. (Sólo el primer punto se resuelve en la Guía de estudio.)

2. Describir el significado geométrico de las siguientes asociaciones en coordenadas cilíndricas:

(a) $(r, \theta, z) \mapsto (r, \theta, -z)$

(b) $(r, \theta, z) \mapsto (r, \theta + \pi, -z)$

(c) $(r, \theta, z) \mapsto (-r, \theta - \pi/4, z)$

3. Describir el significado geométrico de las siguientes asociaciones en coordenadas esféricas:

(a) $(\rho, \theta, \phi) \mapsto (\rho, \theta + \pi, \phi)$

(b) $(\rho, \theta, \phi) \mapsto (\rho, \theta, \pi - \phi)$

(c) $(\rho, \theta, \phi) \mapsto (2\rho, \theta + \pi/2, \phi)$

4. (a) Describir las superficies $r =$ constante, $\theta =$ constante y $z =$ constante en el sistema de coordenadas cilíndricas.

(b) Describir las superficies $\rho =$ constante, $\theta =$ constante y $\phi =$ constante en el sistema de coordenadas esféricas.

5. Mostrar que para representar cada punto en $\mathbf{R}^3$ por medio de coordenadas esféricas es necesario tomar sólo valores de θ entre 0 y 2π, valores de ϕ entre 0 y π, y valores de $\rho \geq 0$. ¿Serían únicas estas coordenadas si permitimos que $\rho \leq 0$?

6. Usando coordenadas cilíndricas y los vectores ortonormales (ortogonales normalizados) $\mathbf{e}_r$, $\mathbf{e}_\theta$ y $\mathbf{e}_z$ (ver la figura 1.4.8),

(a) expresar cada uno de los vectores $\mathbf{e}_r$, $\mathbf{e}_\theta$ y $\mathbf{e}_z$ en términos de $\mathbf{i}, \mathbf{j}, \mathbf{k}$ y (x, y, z); y

(b) calcular $\mathbf{e}_\theta \times \mathbf{j}$ de manera analítica, usando la parte (a), y geométrica.

7. Usando coordenadas esféricas y los vectores ortonormales (ortogonales normalizados) $\mathbf{e}_\rho$, $\mathbf{e}_\theta$ y $\mathbf{e}_\phi$ (ver la figura 1.4.9),

(a) expresar cada uno de los vectores $\mathbf{e}_\rho$, $\mathbf{e}_\theta$ y $\mathbf{e}_\phi$ en términos de $\mathbf{i}, \mathbf{j}, \mathbf{k}$ y (x, y, z); y

(b) calcular $\mathbf{e}_\theta \times \mathbf{j}$ y $\mathbf{e}_\phi \times \mathbf{j}$ de manera analítica y geométrica.

8. Expresar el plano $z = x$ en coordenadas (a) cilíndricas y (b) esféricas.

9. Mostrar que en coordenadas esféricas:

(a) ρ es la longitud de $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

(b) $\phi = \cos^{-1}(\mathbf{v} \cdot \mathbf{k}/\|\mathbf{v}\|)$, donde $\mathbf{v} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

(c) $\theta = \cos^{-1}(\mathbf{u} \cdot \mathbf{i}/\|\mathbf{u}\|)$, donde $\mathbf{u} = x\mathbf{i} + y\mathbf{j}$.

10. Dos superficies se describen en coordenadas esféricas mediante las ecuaciones $\rho = f(\theta, \phi)$ y $\rho = -2f(\theta, \phi)$, donde $f(\theta, \phi)$ es una función de dos variables. ¿Cómo se obtiene geométricamente la segunda superficie a partir de la primera?

11. Una membrana circular en el espacio está sobre la región $x^2 + y^2 \leq a^2$. La componente z máxima de los puntos de la membrana es b. Suponer que (x, y, z) es un punto en la membrana torcida. Mostrar que el punto correspondiente (r, θ, z) en coordenadas cilíndricas satisface las condiciones $0 \leq r \leq a$, $0 \leq \theta \leq 2\pi$, $|z| \leq b$.

12. Un tanque con forma de cilindro circular recto, con radio de 3 m y altura de 5 m contiene la mitad de líquido y reposa de lado. Describir el espacio ocupado por el aire dentro del tanque mediante la elección adecuada de coordenadas cilíndricas.

13. Se va a diseñar un vibrómetro que soporte los efectos del calentamiento de su cubierta esférica de diámetro d, la cual está enterrada a profundidad de $d/3$; el Sol calienta la parte superior (suponer que la Tierra es plana). El análisis de la conducción de calor requiere una descripción de la parte enterrada de la cubierta, en coordenadas esféricas. Hallarla.

14. Un cartucho de filtro de aceite es un cilindro poroso, circular, recto, por el cual fluye el aceite desde el eje hacia la superficie curvada exterior. Describir el cartucho en coordenadas cilíndricas, si el diámetro del filtro es de 4.5 pulg, la altura es de 5.6 pulg y el centro del cartucho está taladrado (a todo lo largo) desde arriba para permitir la entrada de un perno de $\frac{5}{8}$ de pulg.

***15.** Describir la superficie dada en coordenadas esféricas por $\rho = \cos 2\theta$.

1.5 ESPACIO EUCLIDIANO n-DIMENSIONAL

En las secciones 1.1 y 1.2 estudiamos los espacios $\mathbf{R} = \mathbf{R}^1$, $\mathbf{R}^2$ y $\mathbf{R}^3$ y dimos sus interpretaciones geométricas. Por ejemplo, un punto (x, y, z) en $\mathbf{R}^3$ se puede pensar como un objeto geométrico, a saber, el segmento de recta dirigido, o vector, que sale del origen y termina en el punto (x, y, z). Entonces podemos pensar $\mathbf{R}^3$ de cualquiera de estas dos maneras:

(i) Algebraicamente, como un conjunto de ternas (x, y, z) donde x, y y z son números reales

(ii) Geométricamente, como un conjunto de segmentos dirigidos

Estas dos maneras de considerar $\mathbf{R}^3$ son equivalentes. La definición (i) es más conveniente para generalizar. Específicamente, podemos definir $\mathbf{R}^n$, donde n es un entero positivo (quizá mayor que 3), como el conjunto de todas las n-adas ordenadas $(x_1, x_2, \ldots, x_n)$ donde los x_i son números reales. Por ejemplo, $(1, \sqrt{5}, 2, \sqrt{3}) \in \mathbf{R}^4$.

El conjunto $\mathbf{R}^n$ definido anteriormente se conoce como *n-espacio euclidiano*, y sus elementos $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ se conocen como *vectores* o *n-vectores*. Al hacer $n = 1$, 2 o 3, recuperamos la recta, el plano y el espacio tridimensional, respectivamente.

Comenzamos nuestro estudio del n-espacio euclidiano introduciendo varias operaciones algebraicas. Éstas son análogas a las introducidas en la sección 1.1 para $\mathbf{R}^2$ y $\mathbf{R}^3$. Las primeras dos, suma y multiplicación por un escalar, se definen como sigue:

(i) $(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$;

y

(ii) para cualquier número real α,

$$\alpha(x_1, x_2, \ldots, x_n) = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n).$$

La importancia geométrica de estas operaciones para $\mathbf{R}^2$ y $\mathbf{R}^3$ se analizó en la sección 1.1.

Los n vectores $\mathbf{e}_1 = (1, 0, 0, \ldots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \ldots, 0)$, $\ldots$, $\mathbf{e}_n = (0, 0, \ldots 0, 1)$ se llaman *vectores de la base usual* de $\mathbf{R}^n$, y generalizan los tres vectores unitarios ortogonales entre sí, $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$ de $\mathbf{R}^3$. El vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ se puede escribir como $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$.

Para dos vectores $\mathbf{x} = (x_1, x_2, x_3)$ y $\mathbf{y} = (y_1, y_2, y_3)$ en $\mathbf{R}^3$, definimos el *producto punto* o *producto interno* $\mathbf{x} \cdot \mathbf{y}$ como el número real $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + x_3y_3$. Esta definición se extiende fácilmente a $\mathbf{R}^n$; específicamente, para $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, definimos $\mathbf{x} \cdot \mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n$. En $\mathbf{R}^n$ se suele usar la notación $\langle \mathbf{x}, \mathbf{y} \rangle$ en lugar de $\mathbf{x} \cdot \mathbf{y}$, para el producto interior. Continuando con la analogía con $\mathbf{R}^3$, ahora tenemos que definir el concepto abstracto de *longitud* o *norma* de un vector $\mathbf{x}$ mediante la fórmula

$$\text{longitud de } \mathbf{x} = \|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Si $\mathbf{x}$ y $\mathbf{y}$ son dos vectores en el plano ($\mathbf{R}^2$) o en el espacio ($\mathbf{R}^3$), sabemos que el ángulo θ entre ellos está dado por la fórmula

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \, \|\mathbf{y}\|}.$$

El lado derecho de esta ecuación está definido tanto para $\mathbf{R}^n$ como para $\mathbf{R}^2$. Aún representa el coseno del ángulo entre $\mathbf{x}$ y $\mathbf{y}$; este ángulo está bien definido pues $\mathbf{x}$ y $\mathbf{y}$ están en un subespacio bidimensional de $\mathbf{R}^n$ (el plano determinado por $\mathbf{x}$ y $\mathbf{y}$). El producto punto es una poderosa herramienta matemática; una razón es que incorpora el concepto geométrico de ángulo entre dos vectores.

Será útil disponer de algunas propiedades algebraicas del producto interior. Se resumen en el siguiente teorema (ver las propiedades (i), (ii), (iii) y (iv) de la sección 1.2)

TEOREMA 2 *Para* $\mathbf{x}$, $\mathbf{y}$ *y* $\mathbf{z} \in \mathbf{R}^n$ *y* α *y* β *números reales, tenemos*

(i) $(\alpha\mathbf{x} + \beta\mathbf{y}) \cdot \mathbf{z} = \alpha(\mathbf{x} \cdot \mathbf{z}) + \beta(\mathbf{y} \cdot \mathbf{z})$.

(ii) $\mathbf{x} \cdot \mathbf{y} = \mathbf{y} \cdot \mathbf{x}$.

(iii) $\mathbf{x} \cdot \mathbf{x} \geq 0$.

(iv) $\mathbf{x} \cdot \mathbf{x} = 0$ *si, y sólo si,* $\mathbf{x} = \mathbf{0}$.

DEMOSTRACIÓN Cada una de las cuatro afirmaciones se puede probar mediante un sencillo cálculo. Por ejemplo, para probar la propiedad (i) escribimos

$$\begin{aligned}(\alpha\mathbf{x} + \beta\mathbf{y}) \cdot \mathbf{z} &= (\alpha x_1 + \beta y_1, \alpha x_2 + \beta y_2, \ldots, \alpha x_n + \beta y_n) \cdot (z_1, z_2, \ldots, z_n) \\ &= (\alpha x_1 + \beta y_1)z_1 + (\alpha x_2 + \beta y_2)z_2 + \cdots + (\alpha x_n + \beta y_n)z_n \\ &= \alpha x_1 z_1 + \beta y_1 z_1 + \alpha x_2 z_2 + \beta y_2 z_2 + \cdots + \alpha x_n z_n + \beta y_n z_n \\ &= \alpha(\mathbf{x} \cdot \mathbf{z}) + \beta(\mathbf{y} \cdot \mathbf{z}).\end{aligned}$$

La otra demostración es similar. ■

En la sección 1.2 probamos una propiedad mucho más interesante de los productos punto, llamada la desigualdad de Cauchy-Schwarz (a veces se le llama desigualdad de Cauchy-Bunyakovskii-Schwarz, o simplemente desigualdad CBS, porque se descubrió de manera independiente en casos particulares por el matemático francés Cauchy, el matemático ruso Bunyakovskii y el matemático alemán Schwarz). Para $\mathbf{R}^2$ la demostración requirió de la ley de los cosenos. Podríamos escoger este método para $\mathbf{R}^n$, restringiendo nuestra atención a un plano en $\mathbf{R}^n$. Sin embargo, podemos dar una demostración directa, completamente algebraica.

TEOREMA 3 (DESIGUALDAD DE CAUCHY-SCHWARZ) *Sean* $\mathbf{x}$ *y* $\mathbf{y}$ *vectores en* $\mathbf{R}^n$. *Entonces*

$$|\mathbf{x} \cdot \mathbf{y}| \leq \|\mathbf{x}\| \, \|\mathbf{y}\|.$$

DEMOSTRACIÓN Sea $a = \mathbf{y} \cdot \mathbf{y}$ y $b = -\mathbf{x} \cdot \mathbf{y}$. Si $a = 0$ el teorema es claramente válido, pues entonces $\mathbf{y} = \mathbf{0}$ y ambos lados de la desigualdad se reducen a 0. Así, podemos suponer que $a \neq 0$. Por el teorema 2 tenemos

$$\begin{aligned} 0 \leq (a\mathbf{x} + b\mathbf{y}) \cdot (a\mathbf{x} + b\mathbf{y}) &= a^2 \mathbf{x} \cdot \mathbf{x} + 2ab\mathbf{x} \cdot \mathbf{y} + b^2 \mathbf{y} \cdot \mathbf{y} \\ &= (\mathbf{y} \cdot \mathbf{y})^2 \mathbf{x} \cdot \mathbf{x} - (\mathbf{y} \cdot \mathbf{y})(\mathbf{x} \cdot \mathbf{y})^2. \end{aligned}$$

Al dividir entre $\mathbf{y} \cdot \mathbf{y}$ se tiene

$$0 \leq (\mathbf{y} \cdot \mathbf{y})(\mathbf{x} \cdot \mathbf{x}) - (\mathbf{x} \cdot \mathbf{y})^2$$

o

$$(\mathbf{x} \cdot \mathbf{y})^2 \leq (\mathbf{x} \cdot \mathbf{x})(\mathbf{y} \cdot \mathbf{y}) = \|\mathbf{x}\|^2 \|\mathbf{y}\|^2$$

Al extraer raíz cuadrada en ambos lados de esta desigualdad se obtiene la regla deseada. ■

Hay una consecuencia muy útil de la desigualdad de Cauchy-Schwarz en términos de longitudes. La desigualdad del triángulo es geométricamente clara en $\mathbf{R}^3$. En la figura 1.5.1, $\|\mathrm{OQ}\| = \|\mathbf{x} + \mathbf{y}\|$, $\|\mathrm{OP}\| = \|\mathbf{x}\| = \|\mathrm{RQ}\|$ y $\|\mathrm{OR}\| = \|\mathbf{y}\|$. Como la suma de las longitudes de dos lados de un triángulo es mayor o igual que la longitud del tercero, tenemos $\|\mathrm{OQ}\| \leq \|\mathrm{OR}\| + \|\mathrm{RQ}\|$, esto es $\|\mathbf{x} + \mathbf{y}\| \leq$

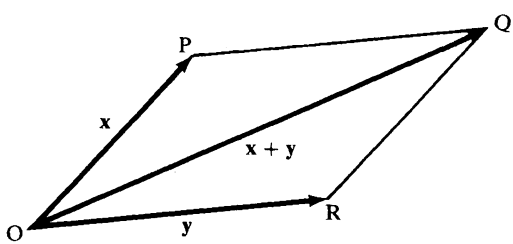


Figura 1.5.1 Esta situación geométrica muestra que $\|\mathrm{OQ}\| \leq \|\mathrm{OR}\| + \|\mathrm{RQ}\|$, o, en notación vectorial, que $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$, lo cual es la desigualdad del triángulo.

$\|\mathbf{x}\|+\|\mathbf{y}\|$. El caso para $\mathbf{R}^n$ no es tan obvio, de modo que daremos la demostración analítica.

COROLARIO *Sean* $\mathbf{x}$ *y* $\mathbf{y}$ *vectores en* $\mathbf{R}^n$. *Entonces*

$$\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \qquad \textit{(desigualdad del triángulo).}$$

DEMOSTRACIÓN Por el teorema 3, $\mathbf{x}\cdot\mathbf{y} \leq |\mathbf{x}\cdot\mathbf{y}| \leq \|\mathbf{x}\|\,\|\mathbf{y}\|$, de modo que

$$\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + 2\mathbf{x}\cdot\mathbf{y} + \|\mathbf{y}\|^2 \leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\|\,\|\mathbf{y}\| + \|\mathbf{y}\|^2.$$

De aquí obtenemos $\|\mathbf{x}+\mathbf{y}\|^2 \leq (\|\mathbf{x}\| + \|\mathbf{y}\|)^2$; al extraer raíz cuadrada se tiene el resultado. ■

Si el teorema 3 y su corolario se desarrollaran algebraicamente, se convertirían en las siguientes útiles desigualdades:

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leq \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \left(\sum_{i=1}^{n} y_i^2\right)^{1/2};$$

$$\left(\sum_{i=1}^{n} (x_i + y_i)^2\right)^{1/2} \leq \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} + \left(\sum_{i=1}^{n} y_i^2\right)^{1/2}$$

EJEMPLO 1 *Sea* $\mathbf{x} = (1, 2, 0, -1)$ *y* $\mathbf{y} = (-1, 1, 1, 0)$. *Verificar el teorema 3 y su corolario para este caso:*

SOLUCIÓN

$$\|\mathbf{x}\| = \sqrt{1^2 + 2^2 + 0^2 + (-1)^2} = \sqrt{6}$$

$$\|\mathbf{y}\| = \sqrt{(-1)^2 + 1^2 + 1^2 + 0^2} = \sqrt{3}$$

$$\mathbf{x}\cdot\mathbf{y} = 1(-1) + 2\cdot 1 + 0\cdot 1 + (-1)0 = 1$$

$$\mathbf{x}+\mathbf{y} = (0, 3, 1, -1)$$

$$\|\mathbf{x}+\mathbf{y}\| = \sqrt{0^2 + 3^2 + 1^2 + (-1)^2} = \sqrt{11}.$$

Calculamos $\mathbf{x}\cdot\mathbf{y} = 1 \leq 4.24 \approx \sqrt{6}\sqrt{3} = \|\mathbf{x}\|\,\|\mathbf{y}\|$, lo cual verifica el teorema 3. De manera análoga podemos verificar su corolario:

$$\|\mathbf{x}+\mathbf{y}\| = \sqrt{11} \approx 3.32 \leq 4.18$$

$$= 2.45 + 1.73 \approx \sqrt{6} + \sqrt{3} = \|\mathbf{x}\| + \|\mathbf{y}\|. \qquad \blacktriangle$$

Por analogía con $\mathbf{R}^3$, podemos definir el concepto de distancia en $\mathbf{R}^n$; a saber, si $\mathbf{x}$ y $\mathbf{y}$ son puntos en $\mathbf{R}^n$, la *distancia entre* $\mathbf{x}$ *y* $\mathbf{y}$ se define como $\|\mathbf{x}-\mathbf{y}\|$, o la longitud del vector $\mathbf{x}-\mathbf{y}$. Insistimos en que *no hay producto cruz* definido en $\mathbf{R}^n$, excepto para $n = 3$. Sólo se generaliza el producto punto.

Generalizando las matrices de 2×2 y de 3×3 (ver la sección 1.3), podemos considerar matrices de $m \times n$, arreglos de mn números:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

También escribiremos A como $[a_{ij}]$. Definimos suma y multiplicación por un escalar por componentes, tal y como se hizo para vectores. Dadas dos matrices de $m \times n$ A y B, podemos sumarlas (restarlas) para obtener una nueva matriz de $m \times n$, $C = A + B$ ($C = A - B$), cuyo ij-ésimo registro c_{ij} es la suma (diferencia) de a_{ij} y b_{ij}. Es claro que $A + B = B + A$.

EJEMPLO 2

(a) $$\begin{bmatrix} 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 1 & 3 \\ 0 & 0 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 8 \end{bmatrix}.$$

(b) $[1 \quad 2] + [0 \quad -1] = [1 \quad 1]$.

(c) $$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$ ▲

Dado un escalar λ y una matriz A de $m \times n$, podemos multiplicar A por λ para obtener una nueva matriz $m \times n$ $\lambda A = C$, cuyo ij-ésimo registro c_{ij} es el producto λa_{ij}.

EJEMPLO 3

$$3 \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & 5 \\ 1 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 3 & -3 & 6 \\ 0 & 3 & 15 \\ 3 & 0 & 9 \end{bmatrix}.$$ ▲

A continuación pasamos a la multiplicación de matrices. Si $A = [a_{ij}]$, $B = [b_{ij}]$, entonces $AB = C$ tiene registros dados por

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj},$$

que es el producto punto del i-ésimo renglón de A y la j-ésima columna de B:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{nn} \end{bmatrix}.$$

EJEMPLO 4 Sea

$$A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

Entonces

$$AB = \begin{bmatrix} 0 & 4 & 3 \\ 1 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{y} \quad BA = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 3 \\ 3 & 1 & 0 \end{bmatrix}. \quad \blacktriangle$$

De manera análoga, mediante la misma regla podemos multiplicar una matriz de $n \times m$ (n renglones, m columnas) por una matriz de $m \times p$ (m renglones, p columnas) para obtener una matriz de $n \times p$ (n renglones, p columnas). Nótese que para que esté definida AB, el número de columnas de A debe ser igual al número de renglones de B.

EJEMPLO 5 Sea

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Entonces

$$AB = \begin{bmatrix} 3 & 1 & 5 \\ 3 & 4 & 5 \end{bmatrix},$$

y BA no está definida. $\blacktriangle$

EJEMPLO 6 Sea

$$A = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix} \quad \text{y} \quad B = [2 \quad 2 \quad 1 \quad 2].$$

Entonces

$$AB = \begin{bmatrix} 2 & 2 & 1 & 2 \\ 4 & 4 & 2 & 4 \\ 2 & 2 & 1 & 2 \\ 6 & 6 & 3 & 6 \end{bmatrix} \quad \text{y} \quad BA = [13]. \quad \blacktriangle$$

Cualquier matriz A de $m \times n$ determina una asociación de $\mathbf{R}^n$ a $\mathbf{R}^m$ de la manera siguiente: Sea $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbf{R}^n$; considerar la matriz columna de

$n \times 1$ asociada con $\mathbf{x}$, que denotaremos *temporalmente* por $\mathbf{x}^T$:

$$\mathbf{x}^T = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

y multiplicar A por $\mathbf{x}^T$ (considerada como una matriz de $n \times 1$) para obtener una nueva matriz de $m \times 1$:

$$A\mathbf{x}^T = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

Entonces obtenemos un vector $\mathbf{y} = (y_1, \ldots, y_m)$. Para usar una matriz A que sirva para obtener una asociación de los vectores $\mathbf{x} = (x_1, \ldots, x_n)$ a vectores $\mathbf{y} = (y_1, \ldots, y_m)$ de acuerdo con la ecuación anterior, hemos de escribir los vectores en forma de columna $\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ en lugar de la forma de renglón $(x_1, \ldots, x_n)$. Este cambio repentino de escribir $\mathbf{x}$ como renglón a escribirlo como columna es necesario debido a las convenciones sobre multiplicación.* Así, aunque cause alguna confusión, escribiremos $\mathbf{x} = (x_1, \ldots, x_n)$ y $\mathbf{y} = (y_1, \ldots, y_m)$ como vectores columna $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$ cuando se trate de multiplicaciones de matrices; esto es, *identificaremos* estas dos formas de escribir vectores. Así, suprimiremos la T en $\mathbf{x}^T$ y consideraremos iguales a $\mathbf{x}^T$ y a $\mathbf{x}$; esto es, $\mathbf{x} = \mathbf{x}^T$.

Así, $A\mathbf{x} = \mathbf{y}$ significará "en realidad" lo siguiente: Escríbase $\mathbf{x}$ como vector columna, multiplíquese por A, y sea $\mathbf{y}$ el vector cuyas componentes son las del vector columna resultante de la multiplicación. La regla $\mathbf{x} \mapsto A\mathbf{x}$ define, por lo tanto, una asociación de $\mathbf{R}^n$ a $\mathbf{R}^m$. Esta asociación es lineal; esto es, satisface

$$A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$$
$$A(\alpha\mathbf{x}) = \alpha(A\mathbf{x}), \qquad \alpha \text{ un escalar},$$

como puede verificarse fácilmente. En un curso de álgebra lineal se aprende que, recíprocamente, cualquier transformación lineal de $\mathbf{R}^n$ a $\mathbf{R}^m$ se puede representar mediante una matriz de $m \times n$.

Si $A = [a_{ij}]$ es una matriz de $m \times n$ y $\mathbf{e}_j$ es el j-ésimo vector de la base usual de $\mathbf{R}^n$, entonces $A\mathbf{e}_j$ es un vector en $\mathbf{R}^m$ con componentes iguales a las de la j-ésima columna de A. Esto es, la i-ésima componente de $A\mathbf{e}_j$ es a_{ij}. En símbolos, $(A\mathbf{e}_j)_i = a_{ij}$.

*Si los matemáticos hubieran elegido la convención de escribir $\mathbf{x}A$ en lugar de $A\mathbf{x}$, o hubiera diferentes reglas para la multiplicación de matrices, se podría conservar a $\mathbf{x}$ como renglón.

EJEMPLO 7 Si

$$A = \begin{bmatrix} 1 & 0 & 3 \\ -1 & 0 & 1 \\ 2 & 1 & 2 \\ -1 & 2 & 1 \end{bmatrix},$$

entonces $\mathbf{x} \mapsto A\mathbf{x}$ de $\mathbf{R}^3$ a $\mathbf{R}^4$ es la asociación definida por

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mapsto \begin{bmatrix} x_1 + 3x_3 \\ -x_1 + x_3 \\ 2x_1 + x_2 + 2x_3 \\ -x_1 + 2x_2 + x_3 \end{bmatrix}. \quad \blacktriangle$$

EJEMPLO 8 A continuación se ilustra lo que sucede a un punto particular cuando se manda mediante una matriz de 4×3:

$$Ae_2 = \begin{bmatrix} 4 & 2 & 9 \\ 3 & 5 & 4 \\ 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 2 \\ 1 \end{bmatrix} = \text{segunda columna de } A. \quad \blacktriangle$$

La multiplicación de matrices no es, en general, *conmutativa:* si A y B son matrices de $n \times n$, entonces por lo general

$$AB \neq BA$$

(ver los ejemplos 4, 5 y 6).

Se dice que una matriz de $n \times n$ es *invertible* si existe alguna matriz B tal que

$$AB = BA = I_n,$$

donde

$$I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

es la matriz identidad de $n \times n$: I_n tiene la propiedad de que $I_n C = CI_n = C$ para cualquier matriz C de $n \times n$. Denotamos B por A^{-1} y la llamamos la *inversa* de A. La inversa, cuando existe, es única.

EJEMPLO 9 Si

$$A = \begin{bmatrix} 2 & 4 & 0 \\ 0 & 2 & 1 \\ 3 & 0 & 2 \end{bmatrix}, \qquad \text{entonces} \qquad A^{-1} = \frac{1}{20}\begin{bmatrix} 4 & -8 & 4 \\ 3 & 4 & -2 \\ -6 & 12 & 4 \end{bmatrix},$$

pues $AA^{-1} = I_3 = A^{-1}A$, como puede verificarse al efectuarse las multiplicaciones de las matrices. ▲

En álgebra lineal se aprenden métodos para calcular inversas; en este libro no se requieren esos métodos. Si A es invertible, es posible resolver la ecuación $A\mathbf{x} = \mathbf{y}$ para el vector $\mathbf{x}$ multiplicando ambos lados por A^{-1} y obtener $\mathbf{x} = A^{-1}\mathbf{y}$.

En la sección 1.3 definimos el determinante de una matriz de 3×3. Esto se puede generalizar por inducción a determinantes de $n \times n$. Ilustramos aquí cómo escribir el determinante de una matriz de 4×4 en términos de los determinantes de matrices de 3×3:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix}$$

$$+ a_{13}\begin{vmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{vmatrix} - a_{14}\begin{vmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{vmatrix}$$

(ver la fórmula (2) de la sección 1.3; los signos se alternan: $+, -, +, -, \ldots$).

Las propiedades básicas de los determinantes de 3×3 que se revisaron en la sección 1.3, mantienen su validez para determinantes de $n \times n$. En particular, nótese el hecho de que si A es una matriz de $n \times n$ y B es la matriz formada al sumar un múltiplo escalar del k-ésimo renglón (o columna) de A al l-ésimo renglón (o, respectivamente, columna) de A, entonces el determinante de A es igual al determinante de B (ver el ejemplo 10 a continuación).

Un teorema básico del álgebra lineal afirma que una matriz A, de $n \times n$ es invertible si y sólo si el determinante de A no es cero. Otra propiedad básica es que $\det(AB) = (\det A)(\det B)$. En este libro no usaremos muchos detalles de álgebra lineal, de modo que dejaremos estas afirmaciones sin demostración.

EJEMPLO 10 *Sea*

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 0 & 2 \end{bmatrix}.$$

Hallar $\det A$. *¿Tiene* A *inversa?*

SOLUCIÓN Al sumar (-1) × la primera columna a la tercera columna, obtenemos

$$\det A = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 2 & 1 & -2 & 1 \\ 1 & 1 & -1 & 2 \end{vmatrix} = 1 \begin{vmatrix} 1 & 0 & 1 \\ 1 & -2 & 1 \\ 1 & -1 & 2 \end{vmatrix}.$$

Al sumar (-1) × la primera columna a la tercera columna de este determinante de 3×3, obtenemos

$$\det A = \begin{vmatrix} 1 & 0 & 0 \\ 1 & -2 & 0 \\ 1 & -1 & 1 \end{vmatrix} = \begin{vmatrix} -2 & 0 \\ -1 & 1 \end{vmatrix} = -2.$$

Así, $\det A = -2 \neq 0$, de modo que A tiene inversa. ▲

Si tenemos tres matrices A, B y C tales que los productos AB y BC están definidos, entonces los productos $(AB)C$ y $A(BC)$ estarán definidos y serán iguales (esto es, la multiplicación de matrices es asociativa). Llamamos a esto *triple producto* de matrices y lo denotamos por ABC.

EJEMPLO 11 Sea

$$A = \begin{bmatrix} 3 \\ 5 \end{bmatrix}, \qquad B = [1 \quad 1], \qquad \text{y} \qquad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Entonces

$$ABC = A(BC) = \begin{bmatrix} 3 \\ 5 \end{bmatrix} [3] = \begin{bmatrix} 9 \\ 15 \end{bmatrix}. \quad ▲$$

EJEMPLO 12

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}. \quad ▲$$

EJERCICIOS

1. Probar las propiedades (ii) a (iv) expresadas en el teorema 2.

2. Mostrar en $\mathbf{R}^n$ que

(a) $2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2 = \|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2$ (Esto se conoce como la ley del paralelogramo).

(b) $\|\mathbf{x}-\mathbf{y}\|\,\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$

(c) $4\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2$ (A esto se le llama la identidad de polarización.)

Interpretar estos resultados geométricamente en términos del paralelogramo formado por $\mathbf{x}$ y $\mathbf{y}$.

Verificar que se cumplen las desigualdades CBS y del triángulo, para los vectores en los ejercicios 3, 4 y 5:

3. $\mathbf{x} = (2, 0, -1)$, $\mathbf{y} = (4, 0, -2)$

4. $\mathbf{x} = (1, 0, 2, 6)$, $\mathbf{y} = (3, 8, 4, 1)$

5. $\mathbf{x} = (1, -1, 1, -1, 1)$, $\mathbf{y} = (3, 0, 0, 0, 2)$

6. Verificar que si A es una matriz de $n \times n$, la asociación $\mathbf{x} \mapsto A\mathbf{x}$ de $\mathbf{R}^n$ a $\mathbf{R}^n$ es lineal.

7. Calcular AB, $\det A$, $\det B$, $\det(AB)$ y $\det(A + B)$ para

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 3 & 2 \\ 3 & 1 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} -2 & 0 & 2 \\ -1 & 1 & -1 \\ 1 & 4 & 3 \end{bmatrix}.$$

8. Calcular AB, $\det A$, $\det B$, $\det(AB)$ y $\det(A + B)$ para

$$A = \begin{bmatrix} 3 & 0 & 1 \\ 1 & 2 & -1 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

9. Usar inducción en k para probar que si $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbf{R}^n$, entonces

$$\|\mathbf{x}_1 + \cdots + \mathbf{x}_k\| \leq \|\mathbf{x}_1\| + \cdots + \|\mathbf{x}_k\|.$$

10. Probar mediante álgebra lineal, la *identidad de Lagrange*: Para números reales $x_1, \ldots x_n$ y $y_1, \ldots, y_n$,

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 = \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right) - \sum_{i<j}(x_i y_j - x_j y_i)^2.$$

Usar esto para dar otra demostración de la desigualdad de Cauchy-Schwarz en $\mathbf{R}^n$.

***11.** Probar que si A es una matriz de $n \times n$, entonces

(a) $\det(\lambda A) = \lambda^n \det A$; y

(b) si B es una matriz obtenida a partir de A al multiplicar cualquier renglón o columna por un escalar λ, entonces $\det B = \lambda \det A$.

En los ejercicios 12, 13 y 14, A, B y C denotan matrices de $n \times n$.

12. ¿Es cierto que $\det(A + B) = \det A + \det B$? Dar una demostración o un contraejemplo.

13. ¿Es cierto que $(A + B)(A - B) = A^2 - B^2$?

14. Suponiendo cierta la ley $\det(AB) = (\det A)(\det B)$, probar que $\det(ABC) = (\det A)(\det B)(\det C)$.

***15.** (Este ejercicio supone conocimiento de integración de funciones continuas de una variable.) Nótese que la demostración de la desigualdad CBS (teorema 3) depende sólo de las propiedades del producto interno listadas en el teorema 2. Usar esta observación para establecer la siguiente desigualdad para funciones continuas f, g: $[0,1] \to \mathbf{R}$:

$$\left|\int_0^1 f(x)g(x)\,dx\right| \leq \sqrt{\int_0^1 [f(x)]^2\,dx}\sqrt{\int_0^1 [g(x)]^2\,dx}.$$

Hacer esto

(a) verificando que el espacio de funciones continuas de $[0,1]$ a $\mathbf{R}$ forman un espacio vectorial; esto es, podemos pensar las funciones f y g de manera abstracta como "vectores" que pueden sumarse entre sí, y multiplicarse por escalares.

(b) introduciendo el producto interno de funciones

$$f \cdot g = \int_0^1 f(x)g(x)\,dx$$

y verificando que satisfaga las condiciones (i) a (iv) del teorema 2.

***16.** Definir la transpuesta A^T de una matriz A de $n \times n$ como sigue: el ij-ésimo elemento de A^T es a_{ji} donde a_{ij} es el ij-ésimo registro de A. Mostrar que A^T está caracterizada por la siguiente propiedad: Para todo $\mathbf{x}$, $\mathbf{y}$ en $\mathbf{R}^n$,

$$(A^T\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y}).$$

17. Verificar que la inversa de

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{es} \quad \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

18. Usar la respuesta al ejercicio 17 para mostrar que la solución del sistema

$$ax + by = e$$
$$cx + dy = f$$

es

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{ad-dc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} e \\ f \end{bmatrix}.$$

19. Suponiendo cierta la ley $\det(AB) = (\det A)(\det B)$, verificar que $(\det A)(\det A^{-1}) = 1$ y concluir que si A tiene inversa, entonces $\det A \neq 0$.

EJERCICIOS DE REPASO DEL CAPÍTULO 1

1. Sea $\mathbf{v} = 3\mathbf{i} + 4\mathbf{j} + 5\mathbf{k}$ y $\mathbf{w} = \mathbf{i} - \mathbf{j} + \mathbf{k}$. Calcular $\mathbf{v} + \mathbf{w}$, $3\mathbf{v}$, $6\mathbf{v} + 8\mathbf{w}$, $-2\mathbf{v}$, $\mathbf{v} \cdot \mathbf{w}$, $\mathbf{v} \times \mathbf{w}$. Interpretar cada operación geométricamente graficando los vectores.

2. Repetir el ejercicio 1 con $\mathbf{v} = 2\mathbf{j} + \mathbf{k}$ y $\mathbf{w} = -\mathbf{i} - \mathbf{k}$.

3. (a) Hallar la ecuación de la recta que pasa por $(-1, 2, -1)$ en la dirección de $\mathbf{j}$.
(b) Hallar la ecuación de la recta que pasa por $(0, 2, -1)$ y $(-3, 1, 0)$.
(c) Hallar la ecuación del plano perpendicular a $(-2, 1, 2)$ que pasa por $(-1, 1, 3)$.

4. (a) Hallar la ecuación de la recta que pasa por $(0, 1, 0)$ en la dirección de $3\mathbf{i} + \mathbf{k}$.
(b) Hallar la ecuación de la recta que pasa por $(0, 1, 1)$ y $(0, 1, 0)$.
(c) Hallar la ecuación para el plano perpendicular a $(-1, 1, -1)$ que pasa por $(1, 1, 1)$.

5. Calcular $\mathbf{v} \cdot \mathbf{w}$ para el siguiente conjunto de vectores:
(a) $\mathbf{v} = -\mathbf{i} + \mathbf{j}$; $\mathbf{w} = \mathbf{k}$.
(b) $\mathbf{v} = \mathbf{i} + 2\mathbf{j} - \mathbf{k}$; $\mathbf{w} = 3\mathbf{i} + \mathbf{j}$.
(c) $\mathbf{v} = -2\mathbf{i} - \mathbf{j} + \mathbf{k}$; $\mathbf{w} = 3\mathbf{i} + 2\mathbf{j} - 2\mathbf{k}$.

6. Calcular $\mathbf{v} \times \mathbf{w}$ para los vectores del ejercicio 5. (Sólo la parte (b) está resuelta en la Guía de estudio.)

7. Hallar el coseno del ángulo entre los vectores en el ejercicio 5. (Sólo la parte (b) está resuelta en la Guía de estudio.)

8. Hallar el área del paralelogramo generado por los vectores del ejercicio 5. (Sólo la parte (b) está resuelta en la Guía de estudio.)

9. Usar notación vectorial para describir el triángulo en el espacio cuyos vértices son el origen y los extremos de los vectores $\mathbf{a}$ y $\mathbf{b}$.

10. Mostrar que los tres vectores $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ están en el mismo plano que pasa por el origen si y sólo si existen tres escalares α, β y γ, no todos iguales a cero, tales que $\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} = \mathbf{0}$.

11. Para los números reales a_1, a_2, a_3, b_1, b_2 y b_3, mostrar que

$$(a_1b_1 + a_2b_2 + a_3b_3)^2 \leq (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2).$$

12. Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores unitarios que son ortogonales entre sí. Si $\mathbf{a} = \alpha\mathbf{u} + \beta\mathbf{v} + \gamma\mathbf{w}$, mostrar que

$$\alpha = \mathbf{a} \cdot \mathbf{u}, \qquad \beta = \mathbf{a} \cdot \mathbf{v}, \qquad \gamma = \mathbf{a} \cdot \mathbf{w}.$$

Interpretar geométricamente los resultados.

13. Sean $\mathbf{a}$ y $\mathbf{b}$ dos vectores en el plano, $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2)$, y sea λ un número real. Mostrar que el área del paralelogramo determinado por $\mathbf{a}$ y $\mathbf{b} + \lambda\mathbf{a}$ es la misma que la del determinado por $\mathbf{a}$ y $\mathbf{b}$. Esbozar. Relacionar este resultado con una conocida propiedad de los determinantes.

14. Hallar el volumen del paralelepípedo determinado por los vértices $(0, 1, 0)$, $(1, 1, 1)$, $(0, 2, 0)$ y $(3, 1, 2)$.

15. Dados vectores distintos de cero $\mathbf{a}$ y $\mathbf{b}$ en $\mathbf{R}^3$, mostrar que el vector $\mathbf{v} = \|\mathbf{a}\|\mathbf{b} + \|\mathbf{b}\|\mathbf{a}$ biseca el ángulo entre $\mathbf{a}$ y $\mathbf{b}$.

16. Usar métodos vectoriales para probar que la distancia del punto (x_1, y_1) a la recta $ax + by = c$ es

$$\frac{|ax_1 + by_1 - c|}{\sqrt{a^2 + b^2}}.$$

17. Verificar que la dirección de $\mathbf{b} \times \mathbf{c}$ está dada por la regla de la mano derecha, escogiendo dos vectores $\mathbf{b}$ y $\mathbf{c}$ de entre $\mathbf{i}$, $\mathbf{j}$ y $\mathbf{k}$.

18. (a) Suponer que $\mathbf{a} \cdot \mathbf{b} = \mathbf{a}' \cdot \mathbf{b}$ para todo $\mathbf{b}$. Mostrar que $\mathbf{a} = \mathbf{a}'$.
(b) Suponer que $\mathbf{a} \times \mathbf{b} = \mathbf{a}' \times \mathbf{b}$ para todo $\mathbf{b}$. ¿Es cierto que $\mathbf{a} = \mathbf{a}'$?

19. (a) Usando métodos vectoriales mostrar que la distancia entre dos rectas no paralelas l_1 y l_2 está dada por

$$d = \frac{|(\mathbf{v}_2 - \mathbf{v}_1) \cdot (\mathbf{a}_1 \times \mathbf{a}_2)|}{\|\mathbf{a}_1 \times \mathbf{a}_2\|},$$

donde $\mathbf{v}_1$ y $\mathbf{v}_2$ son dos puntos cualesquiera sobre l_1 y l_2, respectivamente, y $\mathbf{a}_1$ y $\mathbf{a}_2$ son las direcciones de l_1 y l_2. [IDEA: Considerar el plano que pasa por l_2 y es paralelo a l_1. Mostrar que $(\mathbf{a}_1 \times \mathbf{a}_2)/\|\mathbf{a}_1 \times \mathbf{a}_2\|$ es una normal unitaria para este plano; ahora, proyectar $\mathbf{v}_2 - \mathbf{v}_1$ sobre esta dirección normal.]

(b) Hallar la distancia entre la recta l_1 determinada por los puntos $(-1, -1, 1)$ y $(0, 0, 0)$, y la recta l_2 determinada por los puntos $(0, -2, 0)$ y $(2, 0, 5)$.

20. Mostrar que dos planos, dados por las ecuaciones $Ax + By + Cz + D_1 = 0$ y $Ax + By + Cz + D_2 = 0$, son paralelos, y que la distancia entre ellos es

$$\frac{|D_1 - D_2|}{\sqrt{A^2 + B^2 + C^2}}.$$

21. (a) Probar que el área del triángulo en el plano con vértices (x_1, y_1), (x_2, y_2) y (x_3, y_3), es el valor absoluto de

$$\frac{1}{2}\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix}.$$

(b) Hallar el área del triángulo con vértices $(1, 2)$, $(0, 1)$ y $(-1, 1)$.

22. Convertir los siguientes puntos de coordenadas cartesianas a cilíndricas y esféricas. Localizar:

(a) $(0, 3, 4)$ (b) $(-\sqrt{2}, 1, 0)$
(c) $(0, 0, 0)$ (d) $(-1, 0, 1)$
(e) $(-2\sqrt{3}, -2, 3)$

23. Convertir los siguientes puntos de coordenadas cilíndricas a cartesianas y esféricas. Localizar:

(a) $(1, \pi/4, 1)$ (b) $(3, \pi/6, -4)$
(c) $(0, \pi/4, 1)$ (d) $(2, -\pi/2, 1)$
(e) $(-2, -\pi/2, 1)$

24. Convertir los siguientes puntos de coordenadas esféricas a cartesianas y cilíndricas. Localizar:

(a) $(1, \pi/2, \pi)$ (b) $(2, -\pi/2, \pi/6)$
(c) $(0, \pi/8, \pi/35)$ (d) $(2, -\pi/2, -\pi)$
(e) $(-1, \pi, \pi/6)$

25. Reescribir la ecuación $z = x^2 - y^2$ usando coordenadas cilíndricas y esféricas.

26. Usando coordenadas esféricas, mostrar que

$$\phi = \cos^{-1}\left(\frac{\mathbf{u}\cdot\mathbf{k}}{\|\mathbf{u}\|}\right)$$

donde $\mathbf{u} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Interpretar geométricamente.

27. Verificar las desigualdades de Cauchy-Schwarz y del triángulo para

$$\mathbf{x} = (3, 2, 1, 0) \quad \text{y} \quad \mathbf{y} = (1, 1, 1, 2).$$

28. Multiplicar las matrices

$$A = \begin{bmatrix} 3 & 0 & 1 \\ 2 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

¿Es cierto que $AB = BA$?

29. (a) Mostrar que para dos matrices A y B, de $n \times n$, y $\mathbf{x} \in \mathbf{R}^n$,

$$(AB)\mathbf{x} = A(B\mathbf{x}).$$

(b) ¿Qué implica la igualdad de la parte (a) respecto a la relación entre la composición de las asociaciones $\mathbf{x} \mapsto B\mathbf{x}$, $\mathbf{y} \mapsto A\mathbf{y}$ y la multiplicación de matrices?

30. Hallar el volumen del paralelepípedo generado por los vectores

$$(1, 0, 1), \quad (1, 1, 1), \quad \text{y} \quad (-3, 2, 0).$$

31. Verificar que cualquier asociación lineal de $\mathbf{R}^n$ a $\mathbf{R}^n$ está determinada por una matriz de $n \times n$; esto es, proviene de una matriz de $n \times n$ de la manera explicada en la pág. 62.

32. Hallar una ecuación para el plano que contiene $(3, -1, 2)$ y la recta $\mathbf{v} = (2, -1, 0) + t(2, 3, 0)$.

33. El trabajo W realizado al mover un objeto de $(0,0)$ a $(7,2)$ sujeto a una fuerza $\mathbf{F}$ es $W = \mathbf{F} \cdot \mathbf{r}$ donde $\mathbf{r}$ es el vector con cabeza en $(7,2)$ y cola en $(0,0)$. Las unidades son metros y kilos.

(a) Suponer que la fuerza $\mathbf{F} = 10\cos\theta\mathbf{i} + 10\,\text{sen}\,\theta\mathbf{j}$. Hallar W en términos de θ.

(b) Suponer que la fuerza $\mathbf{F}$ tiene magnitud de 6 kg y forma un ángulo de $\pi/6$ rad con la horizontal, apuntando a la derecha. Hallar W en metros-kilos.

34. Si una partícula con masa m se mueve a una velocidad $\mathbf{v}$, su *momento* es $\mathbf{p} = m\mathbf{v}$. En un juego de canicas, se tira una canica con masa de 2 gramos (g) con una velocidad de 2 metros por segundo (m/s), la cual choca con dos canicas con masa de 1 g cada una, y se para. Una de las canicas sale con una velocidad de 3 m/s formando un ángulo de 45° con la dirección de incidencia de la canica grande, como en la figura 1.R.1. Suponiendo que el momento total antes y después de la colisión es el mismo (de acuerdo con la ley de conservación del momento), ¿a qué ángulo y velocidad se movió la segunda canica?

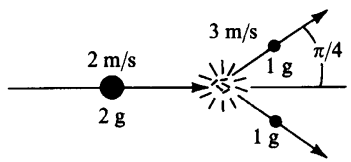


Figura 1.R.1 Momento y canicas.

35. Mostrar que para todo x, y y z,

$$\begin{vmatrix} x+2 & y & z \\ z & y+1 & 10 \\ 5 & 5 & 2 \end{vmatrix} = - \begin{vmatrix} y & x+2 & z \\ 1 & z-x-2 & 10-z \\ 5 & 5 & 2 \end{vmatrix}.$$

36. Mostrar que

$$\begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} \neq 0$$

si x, y y z son todos diferentes.

37. Mostrar que

$$\begin{vmatrix} 66 & 628 & 246 \\ 88 & 435 & 24 \\ 2 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 68 & 627 & 247 \\ 86 & 436 & 23 \\ 2 & -1 & 1 \end{vmatrix}.$$

38. Mostrar que

$$\begin{vmatrix} n & n+1 & n+2 \\ n+3 & n+4 & n+5 \\ n+6 & n+7 & n+8 \end{vmatrix}$$

tiene el mismo valor, sin importar cuánto sea n. ¿Cuál es ese valor?

39. El volumen de un tetraedro con aristas concurrentes $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$ está dado por $V = \frac{1}{6}\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$.

(a) Expresar el volumen como un determinante.

(b) Evaluar V cuando $\mathbf{a} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{b} = \mathbf{i} - \mathbf{j} + \mathbf{k}$, $\mathbf{c} = \mathbf{i} + \mathbf{j}$.

Usar la siguiente definición para los problemas 40 y 41: Sean $\mathbf{r}_1, \ldots, \mathbf{r}_n$ vectores de O a las masas $m_1, \ldots, m_n$. El ***centro de masa*** *es el vector*

$$\mathbf{c} = \left(\sum_{i=1}^{n} m_i \mathbf{r}_i\right) \Big/ \left(\sum_{i=1}^{n} m_i\right).$$

40. Un tetraedro está dado en coordenadas xyz con un vértice en $(0, 0, 0)$ y las tres aristas concurrentes a $(0, 0, 0)$ coinciden con los vectores $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{c}$.

(a) Trazar una figura y rotular las cabezas de los vectores $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$.

(b) Hallar el centro de masa de cada uno de las cuatro caras triangulares del tetraedro si se coloca una unidad de masa en cada vértice.

41. Mostrar que para cualquier vector $\mathbf{r}$, el centro de masa de un sistema satisface

$$\sum_{i=1}^{n} m_i \|\mathbf{r} - \mathbf{r}_i\|^2 = \sum_{i=1}^{n} m_i \|\mathbf{r}_i - \mathbf{c}\|^2 + m\|\mathbf{r} - \mathbf{c}\|^2,$$

donde $m = \sum_{i=1}^{n} m_i$ es la masa total del sistema.

En los ejercicios 42 al 47, hallar un vector unitario que tenga la propiedad dada.

42. Paralelo a la recta $x = 3t + 1$, $y = 16t - 2$, $z = -(t + 2)$.

43. Ortogonal al plano $x - 6y + z = 12$.

44. Paralelo a los planos $8x + y + z = 1$ y $x - y - z = 0$.

45. Ortogonal a $\mathbf{i} + 2\mathbf{j} - \mathbf{k}$ y a $\mathbf{k}$.

46. Ortogonal a la recta $x = 2t - 1$, $y = -t - 1$, $z = t + 2$, y al vector $\mathbf{i} - \mathbf{j}$.

47. A un ángulo de $30°$ con $\mathbf{i}$ y formando ángulos iguales con $\mathbf{j}$ y $\mathbf{k}$.

48. Un par de dipolos distan entre sí en r. (Los *dipolos* son idealizaciones de pequeños magnetos con un polo norte y un polo sur infinitesimalmente cercanos entre sí; la potencia del dipolo se describe por un vector llamado su *momento de dipolo*.) La energía potencial magnética P está dada por $P = -\mathbf{m}_1 \cdot \mathbf{B}_2$ (llamado el *potencial de interacción dipolo-dipolo*), donde el primer dipolo tiene momento $\mathbf{m}_1$ en el campo externo $\mathbf{B}_2$ del segundo dipolo. En unidades MKS,

$$\mathbf{B}_2 = \mu_0 \frac{-\mathbf{m}_2 + 3(\mathbf{m}_2 \cdot \mathbf{l})\mathbf{l}}{4\pi r^3},$$

donde $\mathbf{l}$ es un vector unitario, y μ_0 es una constante escalar.

(a) Mostrar que

$$P = \mu_0 \frac{\mathbf{m}_1 \cdot \mathbf{m}_2 - 3(\mathbf{m}_2 \cdot \mathbf{l})(\mathbf{m}_1 \cdot \mathbf{l})}{4\pi r^3}.$$

(b) Hallar P cuando $\mathbf{m}_1$ y $\mathbf{m}_2$ son perpendiculares.

49. Una esfera de radio 10 centímetros (cm) con centro en $(0, 0, 0)$ rota alrededor del eje z con velocidad angular 4 en una dirección tal que la rotación se ve en sentido contrario al que giran las manecillas del reloj desde el eje z positivo.

(a) Hallar el vector de rotación $\boldsymbol{\omega}$ (ver la sección 1.3, ejercicio 32).

(b) Hallar la velocidad $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$ cuando $\mathbf{r} = 5\sqrt{2}(\mathbf{i} - \mathbf{j})$ está en el "ecuador".

2 DIFERENCIACIÓN

Yo me alejo con espanto y horror, de la triste maldad de las funciones que no tienen derivadas.

CHARLES HERMITE, en una carta a Thomas Jan Stieltjes

En este capítulo se amplían los principios del cálculo diferencial para funciones de una variable a funciones de varias variables. Comenzamos en la sección 2.1 con la geometría de las funciones con valores reales y el estudio de las gráficas de estas funciones como un auxiliar para visualizarlas. En la sección 2.2 se dan algunas definiciones básicas que relacionan límites y continuidad. Este tema se trata superficialmente, pues su desarrollo completo requiere de tiempo y madurez matemática, por lo cual es mejor dejarlo para un curso más avanzado.* Afortunadamente para nuestros propósitos, no es necesario un conocimiento completo de las sutilezas del concepto de límite; el estudiante que tenga dificultad con esta sección deberá tenerlo en mente. Sin embargo, añadimos enseguida, el concepto de límite ocupa un lugar central en la definición de derivada, pero no en los cálculos de las derivadas en problemas específicos; como ya sabemos por los cursos de cálculo de una variable. En las secciones 2.3 y 2.4 se trata la definición de derivada y se obtienen algunas reglas básicas del cálculo: a saber, cómo diferenciar una suma, producto, cociente o composición. En la sección 2.5 estudiamos derivadas direccionales y planos tangentes, relacionando estas ideas con las de la sección 2.1, y en la sección 2.6 consideramos algunas propiedades de las derivadas de orden superior. Finalmente la sección 2.7 es optativa, y ahí se dan algunas demostraciones técnicas.

*Ver, por ejemplo, J. Marsden, *Elementary Classical Analysis*, Freeman, Nueva York, 1974.

Al generalizar el cálculo de una a varias dimensiones, suele ser conveniente usar el lenguaje del álgebra de matrices. En la sección 1.5 hemos resumido lo que vamos a necesitar.

2.1 GEOMETRÍA DE LAS FUNCIONES CON VALORES REALES

Iniciamos nuestra investigación de funciones con valores reales desarrollando métodos para visualizarlas. Introduciremos en particular, los conceptos de gráfica, curva de nivel y superficie de nivel de dichas funciones.

Sea f una función cuyo dominio sea un subconjunto A de $\mathbf{R}^n$ y cuya imagen esté contenida en $\mathbf{R}^m$. Con esto queremos decir que a cada $\mathbf{x} = (x_1, \ldots, x_n) \in A$, f asigna un valor $f(\mathbf{x})$, una m-ada en $\mathbf{R}^m$. Dichas funciones f se llaman *funciones con valores vectoriales** si $m > 1$, y *funciones con valores escalares* si $m = 1$. Por ejemplo, la función con valores escalares $f(x, y, z) = (x^2 + y^2 + z^2)^{-3/2}$ manda al conjunto A de $(x, y, z) \neq (0, 0, 0)$ en $\mathbf{R}^3$ ($n = 3$ en este caso) a $\mathbf{R}(m = 1)$. Para denotar f solemos escribir

$$f: (x, y, z) \mapsto (x^2 + y^2 + z^2)^{-3/2}.$$

Nótese que en $\mathbf{R}^3$ solemos usar la notación (x, y, z) en lugar de (x_1, x_2, x_3). En general, la notación $\mathbf{x} \mapsto f(\mathbf{x})$ es útil para indicar el valor al cual se manda un punto $\mathbf{x} \in \mathbf{R}^n$. Escribimos $f: A \subset \mathbf{R}^n \to \mathbf{R}^m$ para expresar que A es el dominio de f (en $\mathbf{R}^n$) y que la imagen está contenida en $\mathbf{R}^m$. También usamos la expresión f manda A *dentro* de $\mathbf{R}^m$. Dichas funciones f se llaman *funciones de varias variables* si $A \subset \mathbf{R}^n$, $n > 1$.

Como otro ejemplo, tomemos la función con valores vectoriales $g: \mathbf{R}^6 \to \mathbf{R}^2$ definida por la regla

$$g(\mathbf{x}) = g(x_1, x_2, x_3, x_4, x_5, x_6) = (x_1 x_2 x_3 x_4 x_5 x_6, \sqrt{x_1^2 + x_6^2}).$$

La primera coordenada del valor de g en $\mathbf{x}$ es el producto de las coordenadas de $\mathbf{x}$.

Las funciones de $\mathbf{R}^n$ a $\mathbf{R}^m$ no son sólo abstracciones matemáticas, sino que surgen de manera natural en problemas estudiados en todas las ciencias. Por ejemplo, para especificar la temperatura T en una región A del espacio se requiere una función $T: A \subset \mathbf{R}^3 \to \mathbf{R}$ ($n = 3, m = 1$); así $T(x, y, z)$ es la temperatura en el punto (x, y, z). Para especificar la velocidad de un fluido moviéndose en el espacio

*Algunos matemáticos escribirían dicha f en negritas, usando la notación $\mathbf{f}(\mathbf{x})$, pues tiene valores vectoriales. Nosotros no lo hacemos así por cuestión de gusto. Usamos negritas principalmente para asociaciones que sean campos vectoriales, los cuales se introducirán más adelante. El concepto de función fue desarrollado durante muchos siglos, ampliándose la definición para cubrir más casos según iban apareciendo. Por ejemplo, en 1667 James Gregory definió una función como "una cantidad obtenida a partir de otras cantidades mediante una sucesión de operaciones algebraicas o mediante cualquier otra operación imaginable." En 1775 Euler dio la siguiente definición: "Si algunas cantidades dependen de otras de manera que varían cuando varían las últimas, entonces se dice que las primeras son función de las últimas."

se requiere una asociación $\mathbf{V}: \mathbf{R}^4 \to \mathbf{R}^3$, donde $\mathbf{V}(x, y, z, t)$ es el vector velocidad del fluido en el punto (x, y, z) del espacio en el tiempo t (ver la figura 2.1.1). Para especificar la tasa de reacción de una solución que consta de seis reactores químicos A, B, C, D, E y F en proporciones x, y, z, w, u y v, se requiere una asociación $\sigma: U \subset \mathbf{R}^6 \to \mathbf{R}$, donde $\sigma(x, y, z, w, u, v)$ da la tasa cuando los químicos están en las proporciones indicadas. Para especificar el vector cardiaco (el vector que indica la magnitud y dirección del flujo de la corriente eléctrica en el corazón) en el tiempo t, se requiere una asociación $\mathbf{c}: \mathbf{R} \to \mathbf{R}^3$, $t \mapsto \mathbf{c}(t)$.

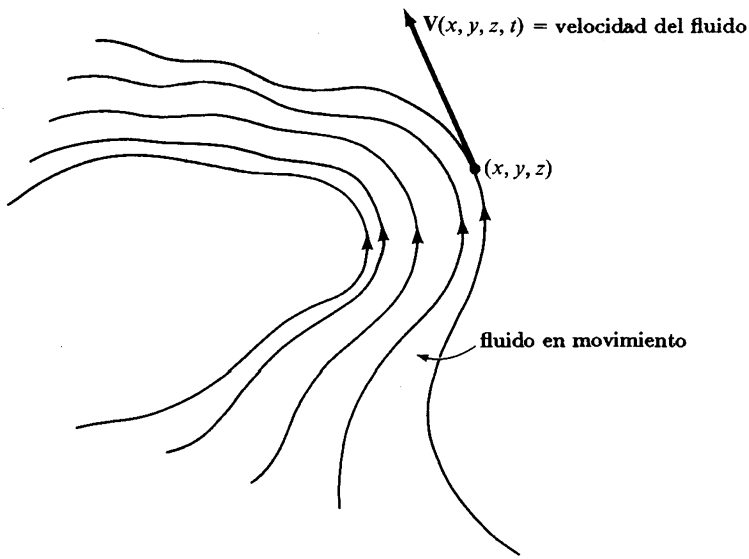


Figura 2.1.1 Un fluido en movimiento define un campo vectorial $\mathbf{V}$ al especificar la velocidad de las partículas del fluido en cada punto en espacio y tiempo.

Cuando $f: U \subset \mathbf{R}^n \to \mathbf{R}$, decimos que f es una *función de n variables, con dominio U y valores reales*. La razón por la que decimos "n variables" es simplemente que consideramos las coordenadas de un punto $\mathbf{x} = (x_1, \ldots, x_n) \in U$ como n variables, y $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ depende de estas variables. Decimos con "valores reales" porque $f(x_1, \ldots, x_n)$ es un número real. Buena parte de nuestro estudio será acerca de funciones con valores reales, por lo que les daremos atención especial.

Para $f: U \subset \mathbf{R} \to \mathbf{R}$, $(n = 1)$, la *gráfica* de f es el subconjunto de $\mathbf{R}^2$ que consta de los puntos $(x, f(x))$ en el plano, para x en U. Este subconjunto se puede pensar como una curva en $\mathbf{R}^2$. Esto se escribe simbólicamente, como

$$\text{gráfica } f = \{(x, f(x)) \in \mathbf{R}^2 | x \in U\},$$

donde las llaves significan "el conjunto de todos" y la barra vertical significa "tal que". Trazar la gráfica de una función de una variable es un recurso útil

para visualizar el comportamiento real de una función. (Ver la figura 2.1.2.) Será conveniente generalizar la idea de gráfica de una función a funciones de varias variables. Esto conduce a la siguiente definición:

DEFINICIÓN *Sea* $f: U \subset \mathbf{R}^n \to \mathbf{R}$. *Definimos la* ***gráfica*** *de* f *como el subconjunto de* $\mathbf{R}^{n+1}$ *que consta de todos los puntos* $(x_1, \ldots, x_n, f(x_1, \ldots, x_n))$ *en* $\mathbf{R}^{n+1}$ *para* $(x_1, \ldots, x_n)$ *en* U. *En símbolos:*

$$\text{gráfica } f = \{(x_1, \ldots, x_n, f(x_1, \ldots, x_n)) \in \mathbf{R}^{n+1} | (x_1, \ldots, x_n) \in U\}.$$

Para el caso $n = 1$, la gráfica es una curva en $\mathbf{R}^2$, mientras que para $n = 2$ es una superficie en $\mathbf{R}^3$ (ver la figura 2.1.2). Para $n = 3$ es difícil visualizar la gráfica, pues como vivimos en un mundo tridimensional, nos es difícil imaginar conjuntos en $\mathbf{R}^4$. Para superar este obstáculo, introducimos la idea de conjunto de nivel.

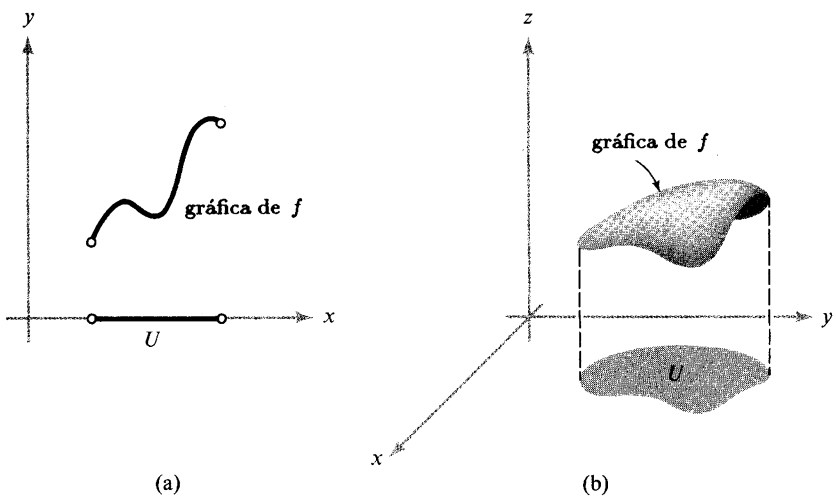


Figura 2.1.2 Gráficas de (a) una función de una variable y (b) una función de dos variables.

Supongamos que $f(x, y, z) = x^2 + y^2 + z^2$. Un *conjunto de nivel* es un subconjunto de $\mathbf{R}^3$ en donde f es constante; por ejemplo, el conjunto donde $x^2+y^2+z^2 = 1$ es un conjunto de nivel para f. A éste sí lo podemos visualizar: es una esfera de radio 1 en $\mathbf{R}^3$. El comportamiento o estructura de una función está determinada en parte por la forma de sus conjuntos de nivel; en consecuencia, entender estos conjuntos nos ayuda a entender la función en cuestión. Los conjuntos de nivel también son útiles para entender funciones de dos variables $f(x, y)$, en cuyo caso hablaremos de *curvas de nivel.*

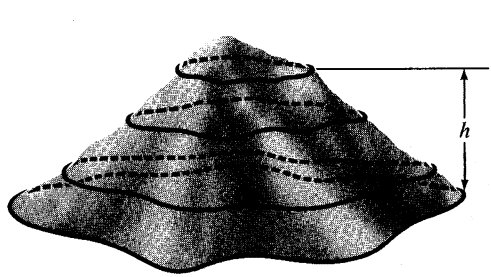


Figura 2.1.3 Los contornos de nivel de una función se definen de la misma manera que las líneas de contorno en un mapa topográfico.

La idea es análoga a la usada para preparar mapas de contornos, donde se trazan líneas para representar altitudes constantes; caminar a lo largo de dicha línea significará caminar en una curva de nivel. En el caso de una colina sobre el plano xy, una gráfica de todas las curvas de nivel nos da una buena idea de la función $h(x, y)$, que representa la altura de la colina en los puntos (x, y) (ver la figura 2.1.3).

DEFINICIÓN *Sea* $f: U \subset \mathbf{R}^n \to \mathbf{R}$ *y sea* $c \in \mathbf{R}$. *Entonces el* ***conjunto de nivel del valor*** c *se define como aquellos puntos* $\mathbf{x} \in U$ *para los cuales* $f(\mathbf{x}) = c$. *Si* $n = 2$, *hablamos de una* ***curva de nivel*** *(de valor c); y si* $n = 3$, *hablamos de una* ***superficie de nivel***. *En símbolos, el conjunto de nivel de valor c se escribe*

$$\{\mathbf{x} \in U | f(\mathbf{x}) = c\} \subset \mathbf{R}^n.$$

Nótese que el conjunto de nivel siempre está en el espacio dominio.

EJEMPLO 1 La función constante $f: \mathbf{R}^2 \to \mathbf{R}$, $(x, y) \mapsto 2$, esto es, la función $f(x, y) = 2$, tiene como gráfica el plano horizontal $z = 2$ en $\mathbf{R}^3$. La curva de nivel del valor c es vacía si $c \neq 2$, y es todo el plano xy si $c = 2$. ▲

EJEMPLO 2 La función $f: \mathbf{R}^2 \to \mathbf{R}$, $(x, y) \mapsto x + y + 2$ tiene como gráfica el plano inclinado $z = x + y + 2$. Este plano insterseca el plano xy ($z = 0$) en la recta $y = -x - 2$ y el eje z en el punto $(0, 0, 2)$. Para cualquier valor $c \in \mathbf{R}$, la curva de nivel del valor c es la recta $y = -x + (c - 2)$; o, en símbolos, el conjunto

$$L_c = \{(x, y) | y = -x + (c - 2)\} \subset \mathbf{R}^2$$

Exhibimos unas cuantas curvas de nivel de la función en la figura 2.1.4. Se trata, en realidad, de un mapa de contorno de la función f. ▲

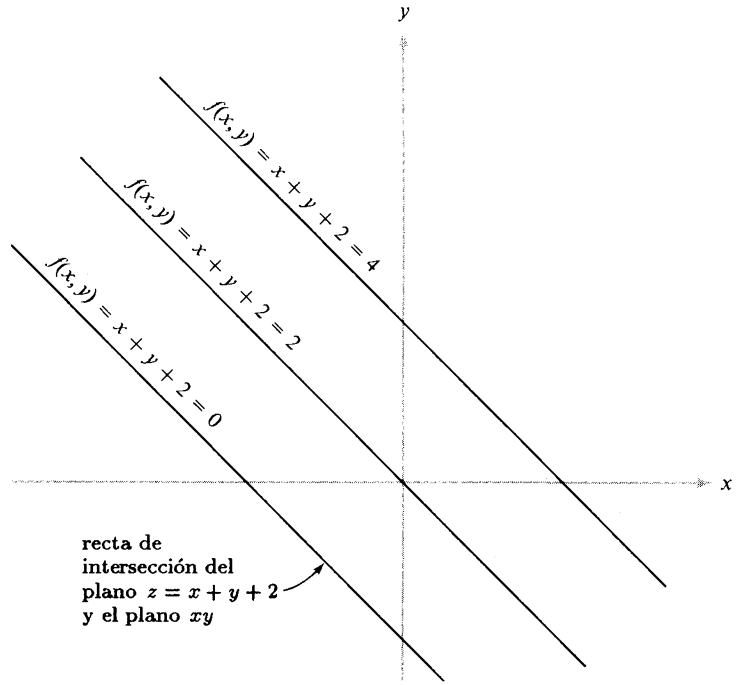


Figura 2.1.4 Las curvas de nivel de $f(x, y) = x + y + z + 2$ muestran el comportamiento de esta función.

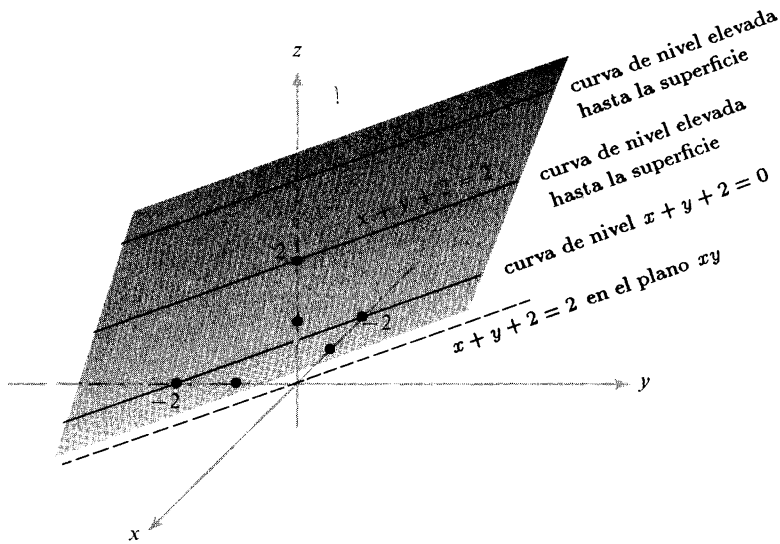


Figura 2.1.5 Relación de las curvas de nivel en la figura 2.1.4, con la gráfica de la función $f(x, y) = x + y + 2$, la cual es el plano $z = x + y + 2$.

A partir de las curvas de nivel rotuladas con el valor o "altura" de la función, se puede inferir la gráfica de la función elevando mentalmente cada curva de nivel a la altura apropiada, sin estirarla, inclinarla o deslizarla. Si se contemplara este procedimiento para todas las curvas de nivel L_c, esto es, para todos los valores $c \in \mathbf{R}$, juntas conformarían toda la gráfica de f, como se indicó en la figura 2.1.5 para el ejemplo 2. Si se visualiza la gráfica sólo para un número finito de curvas de nivel, como suele ser el caso, se produce una especie de modelo de contorno, como en la figura 2.1.4. Sin embargo, si f es una función suave, su gráfica será una superficie suave; entonces, al suavizar mentalmente el modelo de contorno se obtiene una buena idea de la gráfica.

EJEMPLO 3 *Describir la gráfica de la función cuadrática* $f: \mathbf{R}^2 \to \mathbf{R}$, $(x, y) \mapsto x^2 + y^2$.

SOLUCIÓN La gráfica es el *paraboloide de revolución* $z = x^2 + y^2$, orientado hacia arriba desde el origen y alrededor del eje z. La curva de nivel del valor c es vacía para $c < 0$; para $c > 0$ la curva de nivel de valor c es el conjunto $\{(x, y) | x^2 + y^2 = c\}$, un círculo de radio $\sqrt{c}$ con centro en el origen. Así, al elevarlo a la altura c sobre el plano xy, el conjunto de nivel es un círculo de radio $\sqrt{c}$, que indica una forma parabólica (ver las figuras 2.1.6 y 2.1.7). ▲

Es posible determinar el aspecto de una gráfica mediante el *método de las secciones*. Una *sección* de la gráfica de f es la intersección de la gráfica con un plano (vertical). Por ejemplo, si P_1 es el plano xz en $\mathbf{R}^3$, definido por $y = 0$, entonces la sección de f en el ejemplo 3 es el conjunto

$$P_1 \cap \text{gráfica } f = \{(x, y, z) | y = 0, z = x^2\},$$

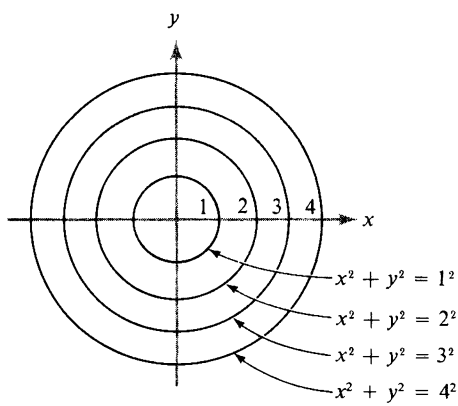


Figura 2.1.6 Algunas curvas de nivel para la función $f(x, y) = x^2 + y^2$.

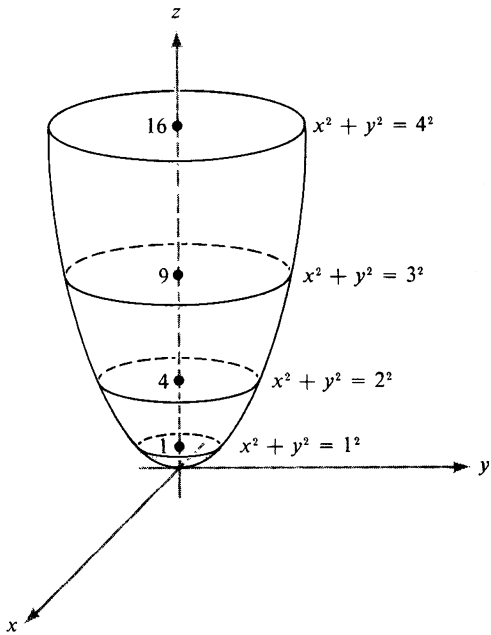


Figura 2.1.7 Las curvas de nivel de la figura 2.1.6 elevadas hasta la gráfica.

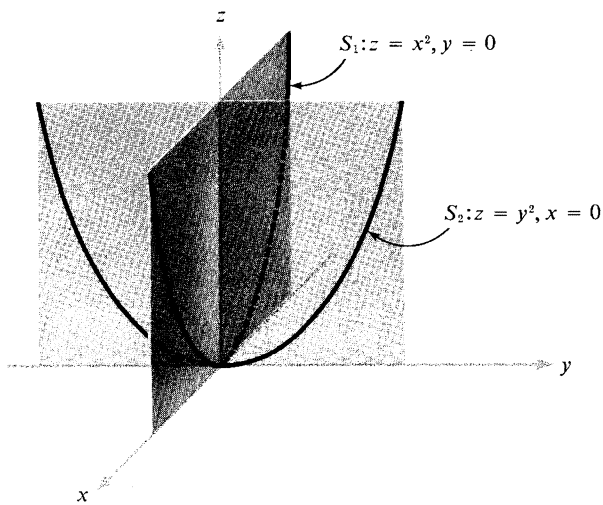


Figura 2.1.8 Dos secciones de la gráfica de $f(x, y) = x^2 + y^2$.

el cual es una parábola en el plano xz. De manera análoga, si P_2 denota al plano yz, definido por $x = 0$, entonces la sección

$$P_2 \cap \text{gráfica } f = \{(x, y, z) | x = 0, z = y^2\}$$

es una parábola en el plano yz (ver la figura 2.1.8). Usualmente, es útil calcular al menos una sección para complementar la información dada por los conjuntos de nivel.

EJEMPLO 4 *La gráfica de la función cuadrática* $f: \mathbf{R}^2 \to \mathbf{R}$, $(x, y) \mapsto x^2 - y^2$ *se llama* ***paraboloide hiperbólico*** *o* ***silla de montar,*** *con centro en el origen. Esbozar la gráfica.*

SOLUCIÓN Para visualizar esta superficie trazamos primero las curvas de nivel. Para determinar las curvas de nivel, resolvemos la ecuación $x^2 - y^2 = c$. Consideremos los valores $c = 0, \pm 1, \pm 4$. Para $c = 0$, tenemos $y^2 = x^2$, o $y = \pm x$, de manera que este conjunto de nivel está formado por dos rectas que pasan por el origen. Para $c = 1$, la curva de nivel es $x^2 - y^2 = 1$, o $y = \pm\sqrt{x^2 - 1}$, que

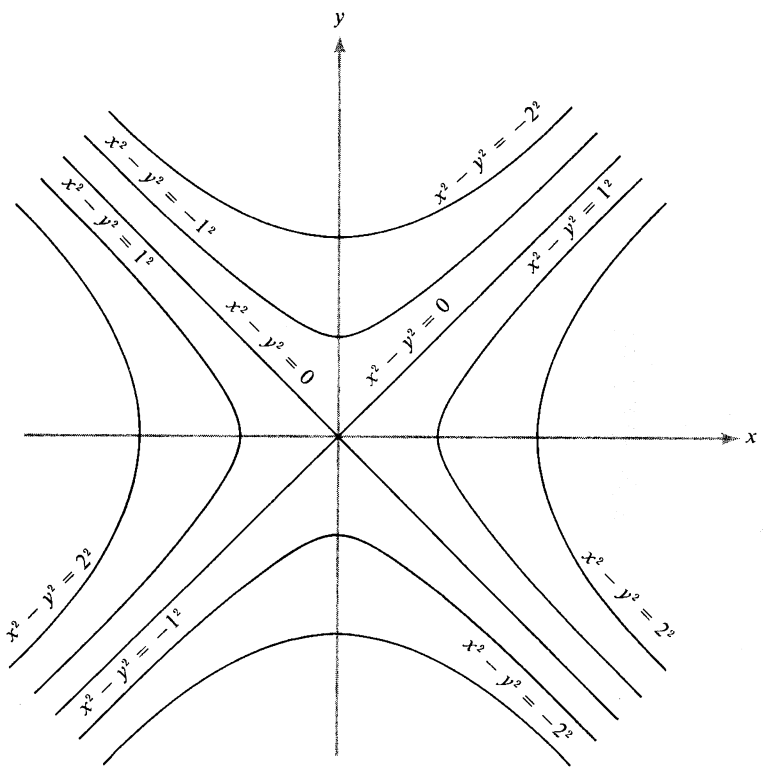


Figura 2.1.9 Curvas de nivel para la función $f(x, y) = x^2 - y^2$.

es una hipérbola que cruza verticalmente el eje x en los puntos $(\pm 1, 0)$ (ver la figura 2.1.9). De manera análoga, para $c = 4$, la curva de nivel está definida por $y = \pm\sqrt{x^2 - 4}$, la hipérbola cruza verticalmente el eje x en $(\pm 2, 0)$. Para $c = -1$, obtenemos la curva $x^2 - y^2 = -1$, esto es, $x = \pm\sqrt{y^2 - 1}$, la hipérbola cruza horizontalmente el eje y en $(0, \pm 1)$. Y para $c = -4$, se obtiene la hipérbola que pasa por $(0, \pm 2)$. Se muestran estas curvas en la figura 2.1.9. Como no es fácil visualizar la gráfica de f a partir sólo de estos datos, calcularemos dos secciones, como lo hicimos en el ejemplo anterior. Para la sección en el plano xz, tenemos

$$P_1 \cap \text{gráfica } f = \{(x, y, z) | y = 0, z = x^2\},$$

que es una parábola abriéndose hacia arriba; y para el plano yz,

$$P_2 \cap \text{gráfica } f = \{(x, y, z) | x = 0, z = -y^2\},$$

que es una parábola abriéndose hacia abajo. Ahora se puede visualizar la gráfica elevando las curvas de nivel a la altura apropiada y suavizando la superficie resultante. Su colocación se facilita al calcular las secciones parabólicas. Este procedimiento genera la silla de montar hiperbólica mostrada en la figura 2.1.10. Comparar esto con las gráficas generadas por computadora en la figura 2.1.11 (nótese que se ha cambiado la orientación de los ejes). ▲

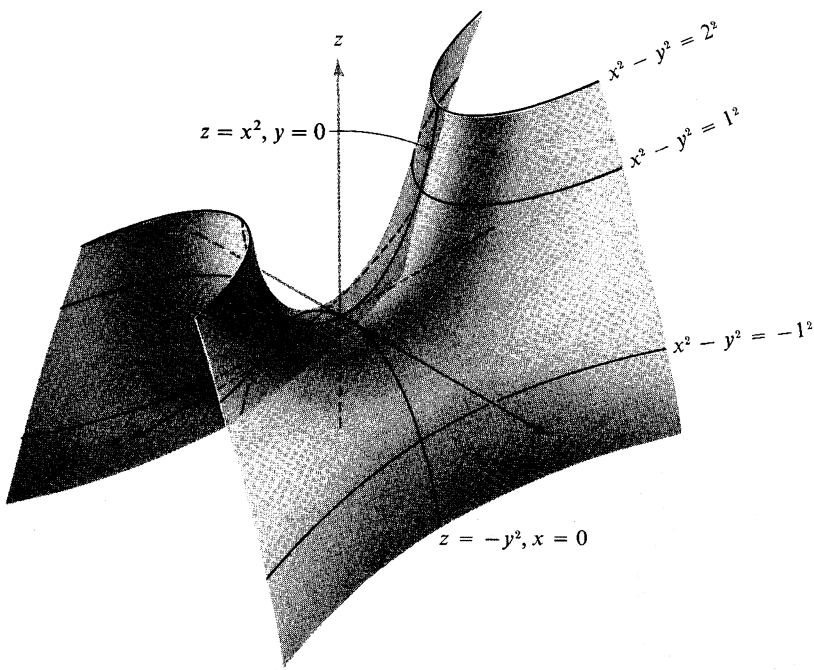


Figura 2.1.10 Algunas curvas de nivel en la gráfica de $f(x, y) = x^2 - y^2$.

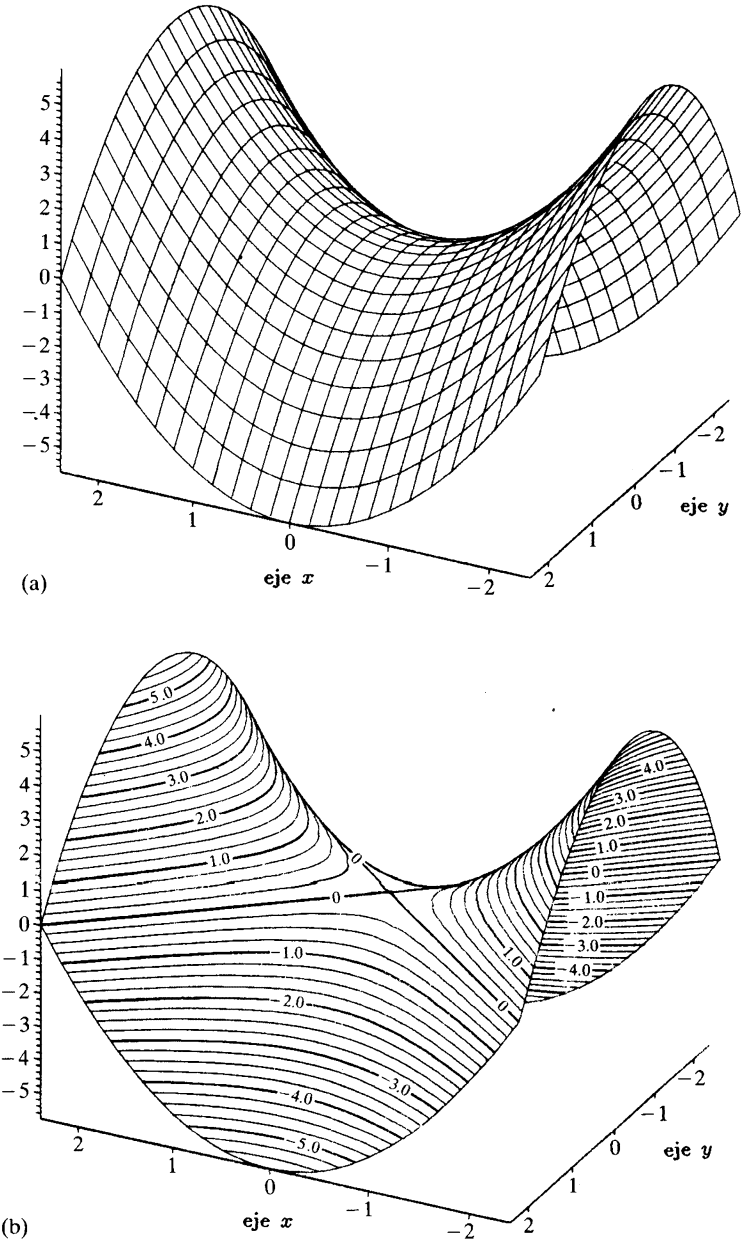


Figura 2.1.11 (a) Gráfica generada por computadora, de $z = x^2 + y^2$. (b) Esta gráfica con las curvas de nivel elevadas.

EJEMPLO 5 *Describir la gráfica de la función* $f: \mathbf{R}^3 \to \mathbf{R}$, $(x, y, z) \mapsto x^2 + y^2 + z^2$.

SOLUCIÓN Éste es el equivalente tridimensional del ejemplo 3. En este contexto, los conjuntos de nivel son superficies en el dominio tridimensional $\mathbf{R}^3$. La gráfica, en $\mathbf{R}^4$, no se puede visualizar directamente; sin embargo se pueden calcular de manera analítica las secciones.

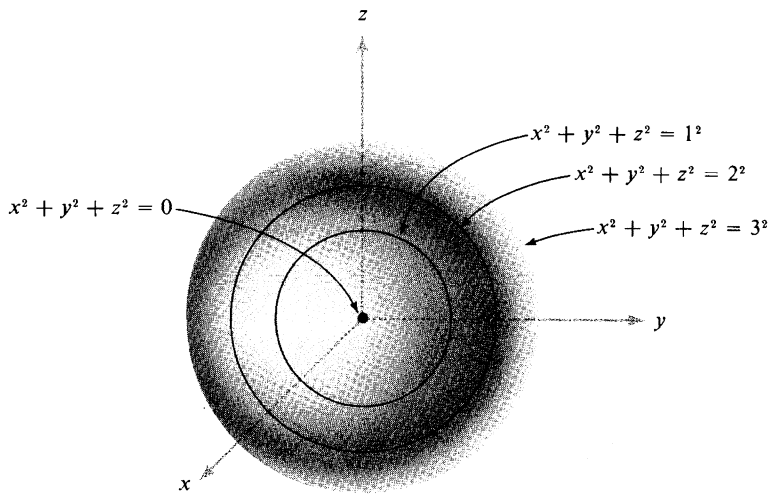


Figura 2.1.12 Algunas superficies de nivel para $f(x, y, z) = x^2 + y^2 + z^2$.

El conjunto de nivel con valor c es el conjunto

$$L_c = \{(x, y, z) | x^2 + y^2 + z^2 = c\},$$

el cual es una esfera con centro en el origen y radio $\sqrt{c}$ para $c > 0$, es un solo punto en el origen para $c = 0$, y es vacío para $c < 0$. En la figura 2.1.12 se muestran los conjuntos de nivel para $c = 0, 1, 4$ y 9. Se obtiene mayor información acerca de la gráfica al calcular una sección. Por ejemplo, si escribimos $S_{z=0} = \{(x, y, z, t) | z = 0\}$, entonces podemos ver la sección

$$S_{z=0} \cap \text{gráfica } f = \{(x, y, z, t) | t = x^2 + y^2, z = 0\}.$$

Como aquí z se mantiene fija en $z = 0$, podemos visualizar esta sección de la gráfica como una superficie en $\mathbf{R}^3$, en las variables x, y y t (figura 2.1.13). La superficie es un paraboloide de revolución. ▲

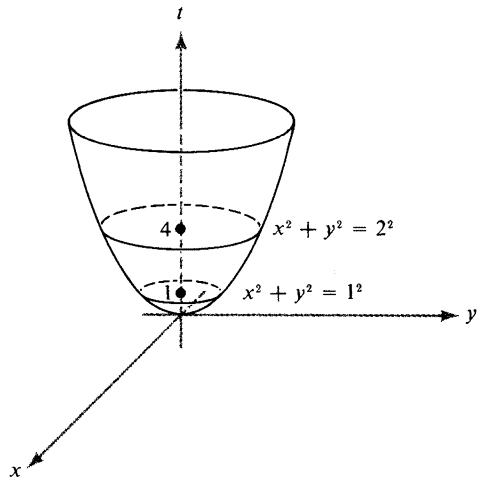


Figura 2.1.13 La sección $z = 0$ de la gráfica de $f(x, y, z) = x^2 + y^2 + z^2$.

EJEMPLO 6 *Describir la gráfica de la función* $f: \mathbf{R}^3 \to \mathbf{R}$ *definida por* $f(x, y, z) = x^2 + y^2 - z^2$, *que es el símil tridimensional del ejemplo 4, y también se conoce como* ***silla de montar***.

SOLUCIÓN Las superficies de nivel están definidas por

$$L_c = \{(x, y, z) | x^2 + y^2 - z^2 = c\}.$$

Para $c = 0$, se trata del cono $z = \pm\sqrt{x^2 + y^2}$ con centro en el eje z. Para c negativa, digamos $c = -a^2$, obtenemos $z = \pm\sqrt{x^2 + y^2 + a^2}$, que es un hiperboloide de dos hojas alrededor del eje z, que atraviesa el eje z en los puntos $(0, 0, \pm a)$. Para c positivo, digamos $c = b^2$, la superficie de nivel es el *hiperboloide de revolución de una hoja* alrededor del eje z definido por $z = \pm\sqrt{x^2 + y^2 - b^2}$, el cual interseca el plano xy en el círculo de radio $|b|$. Estas superficies de nivel se esbozan en la figura 2.1.14. Se puede obtener otra vista de la gráfica a partir de una sección. Por ejemplo, el subespacio $S_{y=0} = \{(x, y, z, t) | y = 0\}$ interseca la gráfica en la sección

$$S_{y=0} \cap \text{gráfica } f = \{(x, y, z, t) | y = 0, t = x^2 - z^2\},$$

esto es, el conjunto de puntos de la forma $(x, 0, z, x^2 - z^2)$, que puede considerarse, como en el ejemplo anterior, una superficie en el espacio xzt (ver la figura 2.1.15). ▲

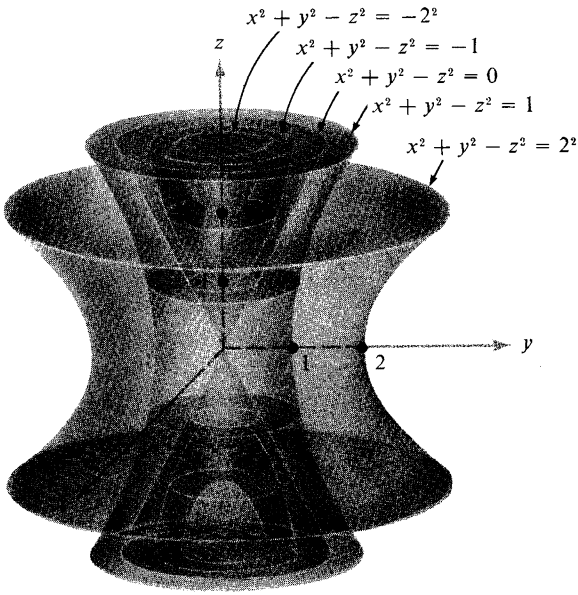


Figura 2.1.14 Algunas superficies de nivel de la función $f(x, y, z) = x^2 + y^2 - z^2$.

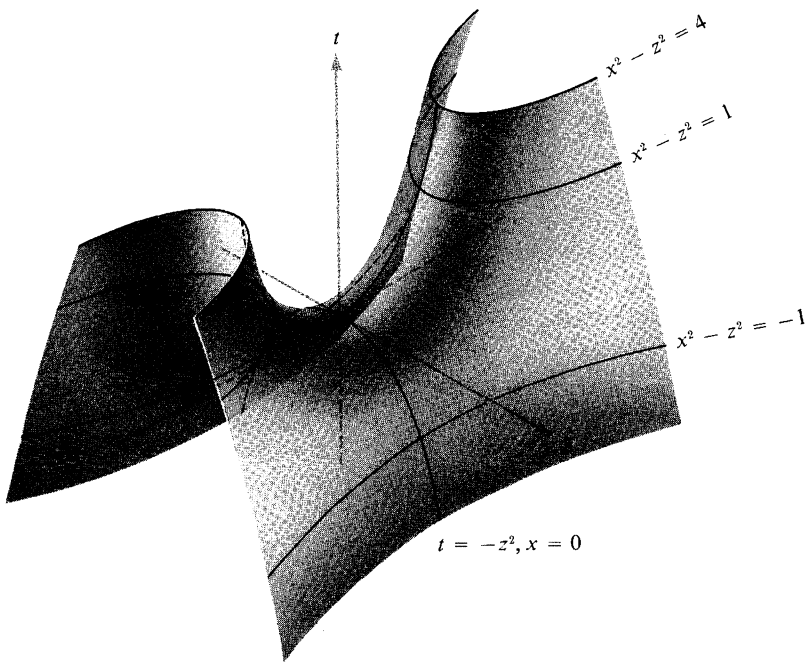
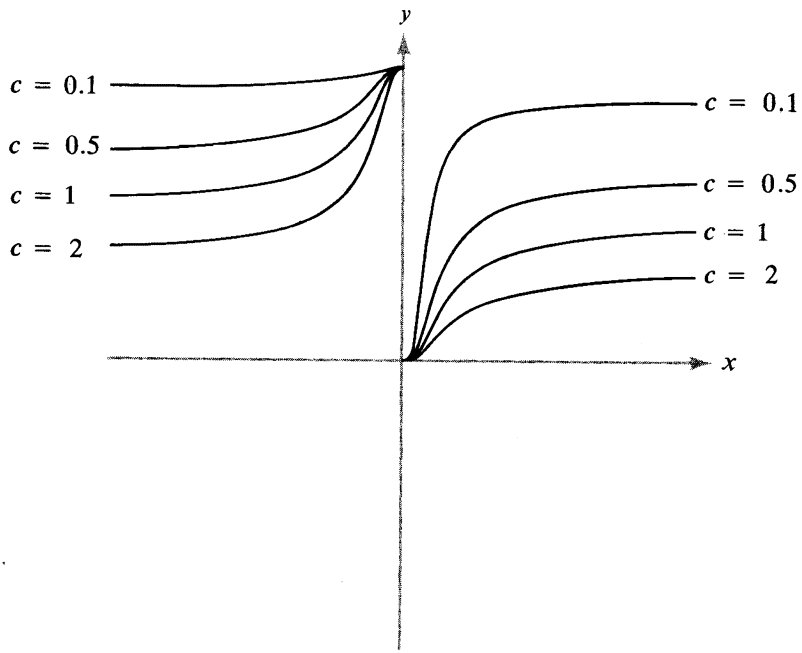


Figura 2.1.15 La sección $y = 0$ de la gráfica de $f(x, y, z) = x^2 + y^2 - z^2$.

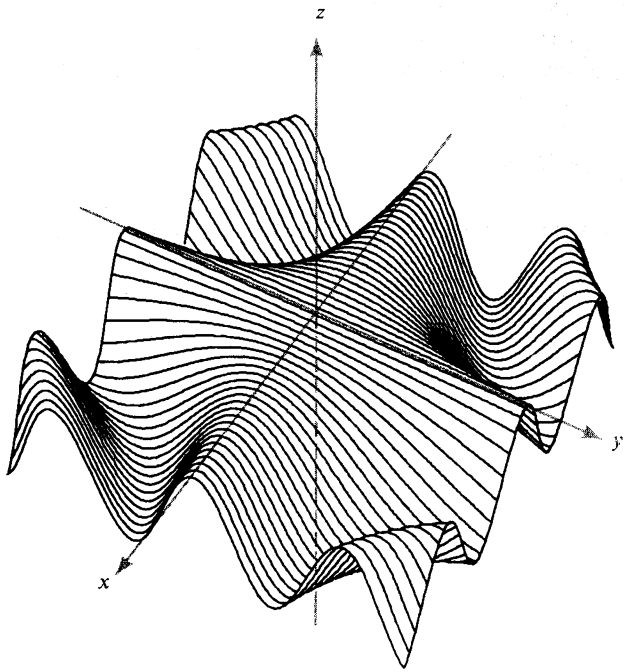
Hemos visto cómo se pueden usar los métodos de secciones y conjuntos de nivel para entender el comportamiento de una función y su gráfica; estas técnicas pueden ser de bastante utilidad para personas que deseen visualizar ampliamente datos complicados.

Existen muchos programas de computadora que pueden trazar una función dada. Para funciones de una variable, se trata sólo de calcular ciertos valores de la función y localizar los puntos. Para funciones de dos variables se usa el método de las secciones. Por ejemplo, para trazar $f(x, y)$, la computadora selecciona secciones paralelas a los ejes, asignando valores, digamos a y y trazando la gráfica correspondiente, después cambiando y y repitiendo el proceso. Así se puede barrer una buena parte de la gráfica. Se dan algunos ejemplos en las figuras 2.1.16 y 2.1.17.

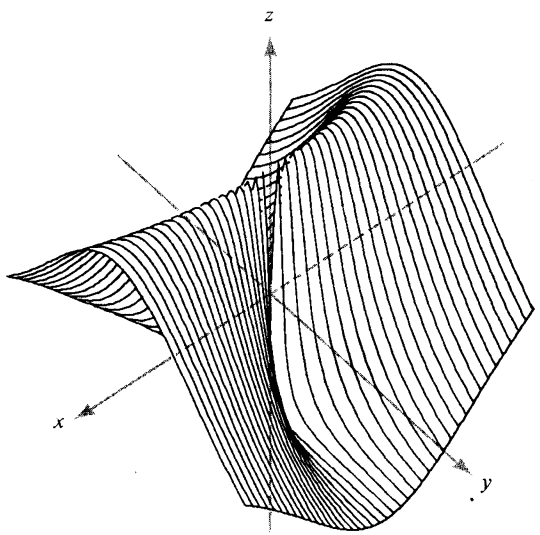


(a) $y = 1/(1 + ce^{1/x})$ *para* $c = 0.1, 0.5, 1, y\ 2$.

Figura 2.1.16 Algunas gráficas generadas por computadora.

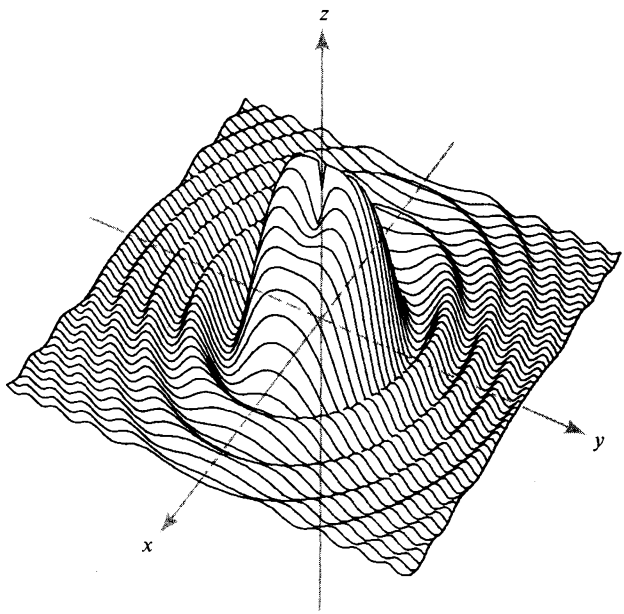


(b) $z = \cos(xy), -3 \leq x \leq 3, -3 \leq y \leq 3.$

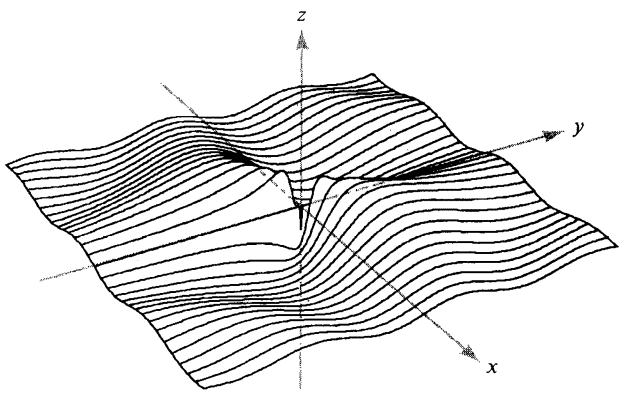


(c) $z = \dfrac{(x^2 - y^2)}{(x^2 + y^2)}, 0 < |x| \leq 1, 0 < |y| \leq 1.$

Figura 2.1.16 Algunas gráficas generadas por computadora. (*Continúa*)



(d) $z = \dfrac{[\text{sen}(2x^2 + 3y^2)]}{[x^2 + y^2]}, 0 < |x| \leq 3, 0 < |y| \leq 3.$



(e) $z = \dfrac{(\text{sen } xy)}{(x^2 + y^2)}, 0 < |x| \leq 3, 0 < |y| \leq 3.$

Figura 2.1.16 (*Continúa*)

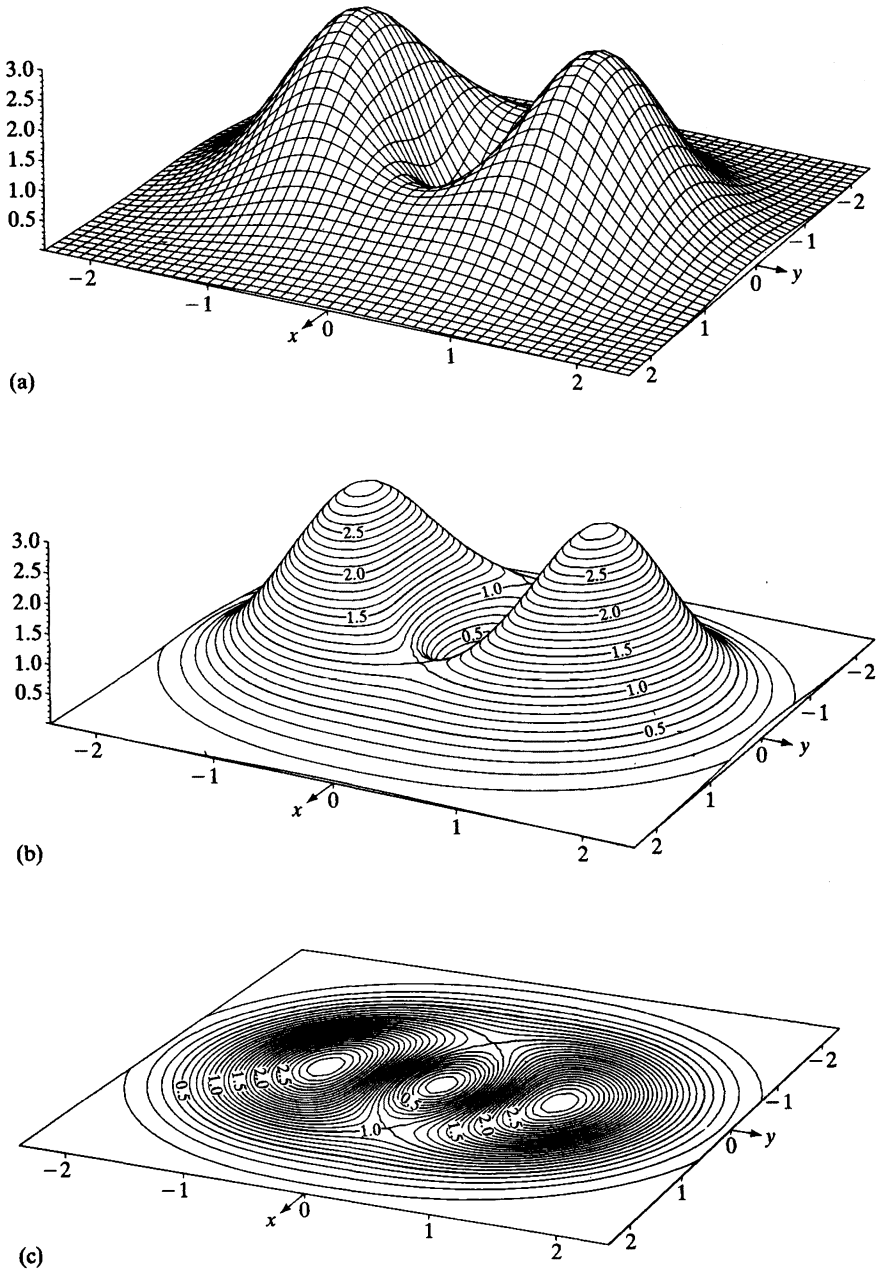


Figura 2.1.17 Gráfica generada por computadora, de $z = (x^2 + 3y^2)\exp(1 - x^2 - y^2)$, representada de cuatro maneras: (a) por secciones, (b) por curvas de nivel sobre la gráfica, (c) por curvas de nivel en el plano xy visto en perspectiva, y (d) por curvas de nivel en el plano xy visto desde arriba.

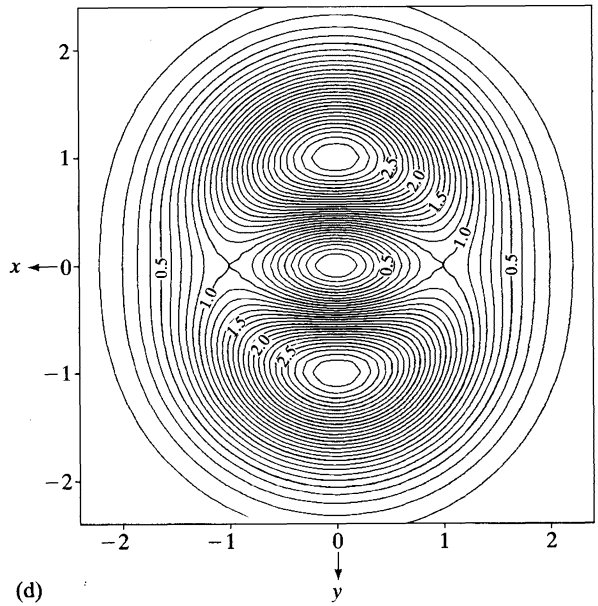


Figura 2.1.17 (*Continúa*)

EJERCICIOS

1. Esbozar curvas de nivel y gráfica de las siguientes funciones:
(a) $f: \mathbf{R}^2 \to \mathbf{R}, (x, y) \mapsto x - y + 2$ (b) $f: \mathbf{R}^2 \to \mathbf{R}, (x, y) \mapsto x^2 + 4y^2$
(c) $f: \mathbf{R}^2 \to \mathbf{R}, (x, y) \mapsto -xy$

2. Describir el comportamiento, conforme varía c, de la curva de nivel $f(x, y) = c$ para cada una de estas funciones:
(a) $f(x, y) = x^2 + y^2 + 1$ (b) $f(x, y) = 1 - x^2 - y^2$ (c) $f(x, y) = x^3 - x$

3. Para las funciones en los ejemplos 2, 3 y 4, calcular la sección de la gráfica definida por el plano

$$S_\theta = \{(x, y, z) | y = x \tan \theta\}$$

para una constante dada θ. Hacer esto expresando z como función de r, donde $x = r \cos \theta$, $y = r \operatorname{sen} \theta$. Determinar cuáles de estas funciones f tienen la propiedad de que la forma de $S_\theta \cap$ gráfica de f sea independiente de θ. (La solución sólo al ejemplo 3 está en la Guía de estudio de este libro.)

En los ejercicios 4 al 10, trazar las curvas de nivel (en el plano xy) para las funciones dadas f y valores especificados de c. Esbozar la gráfica de $z = f(x, y)$.

4. $f(x, y) = 4 - 3x + 2y$, $c = 0, 1, 2, 3, -1, -2, -3$

5. $f(x, y) = (100 - x^2 - y^2)^{1/2}$, $c = 0, 2, 4, 6, 8, 10$

6. $f(x,y) = (x^2 + y^2)^{1/2}$, $c = 0, 1, 2, 3, 4, 5$

7. $f(x,y) = x^2 + y^2$, $c = 0, 1, 2, 3, 4, 5$

8. $f(x,y) = 3x - 7y$, $c = 0, 1, 2, 3, -1, -2, -3$

9. $f(x,y) = x^2 + xy$, $c = 0, 1, 2, 3, -1, -2, -3$

10. $f(x,y) = x/y$, $c = 0, 1, 2, 3, -1, -2, -3$

En los ejercicios 11 al 13, esbozar o describir las superficies de nivel y una sección de la gráfica de cada función.

11. $f: \mathbf{R}^3 \to \mathbf{R}, (x,y,z) \mapsto -x^2 - y^2 - z^2$

12. $f: \mathbf{R}^3 \to \mathbf{R}, (x,y,z) \mapsto 4x^2 + y^2 + 9z^2$

13. $f: \mathbf{R}^3 \to \mathbf{R}, (x,y,z) \mapsto x^2 + y^2$

En los ejercicios 14 al 18, describir la gráfica de cada función calculando algunos conjuntos de nivel y secciones.

14. $f: \mathbf{R}^3 \to \mathbf{R}, (x,y,z) \mapsto xy$

15. $f: \mathbf{R}^3 \to \mathbf{R}, (x,y,z) \mapsto xy + yz$

16. $f: \mathbf{R}^3 \to \mathbf{R}, (x,y,z) \mapsto xy + z^2$

17. $f: \mathbf{R}^2 \to \mathbf{R}, (x,y) \mapsto |y|$

18. $f: \mathbf{R}^2 \to \mathbf{R}, (x,y) \mapsto \text{máx}(|x|, |y|)$

Esbozar o describir las superficies en $\mathbf{R}^3$ *de las ecuaciones presentadas en los ejercicios 19 al 31.*

19. $4x^2 + y^2 = 16$

20. $x + 2z = 4$

21. $z^2 = y^2 + 4$

22. $x^2 + y^2 - 2x = 0$

23. $\dfrac{x}{4} = \dfrac{y^2}{4} + \dfrac{z^2}{9}$

24. $\dfrac{y^2}{9} + \dfrac{z^2}{4} = 1 + \dfrac{x^2}{16}$

25. $z = x^2$

26. $y^2 + z^2 = 4$

27. $z = \dfrac{y^2}{4} - \dfrac{x^2}{9}$

28. $y^2 = x^2 + z^2$

29. $4x^2 - 3y^2 + 2z^2 = 0$

30. $\dfrac{x^2}{9} + \dfrac{y^2}{12} + \dfrac{z^2}{9} = 1$

31. $x^2 + y^2 + z^2 + 4x - by + 9z - b = 0$, donde b es una constante

32. Usando coordenadas polares, describir las curvas de nivel de la función

$$f: \mathbf{R}^2 \to \mathbf{R}, (x, y) \mapsto \begin{cases} \dfrac{2xy}{x^2 + y^2} & \text{si} \quad (x, y) \neq (0, 0) \\ 0 & \text{si} \quad (x, y) = (0, 0) \end{cases}$$

33. Sea $f: \mathbf{R}^2 \backslash \{0\} \to \mathbf{R}$ dada en coordenadas polares por $f(r, \theta) = (\cos 2\theta)/r^2$. Esbozar algunas curvas de nivel respecto a los ejes xy. ($\mathbf{R}^2 \backslash \{0\} = \{\mathbf{x} \in \mathbf{R}^2 | \mathbf{x} \neq \mathbf{0}\}$.)

34. En la figura 2.1.17(d), la "curva" de nivel $z = 3$ aparece formada por dos puntos. Probar esto algebraicamente.

2.2 LÍMITES Y CONTINUIDAD

En esta sección se desarrolla la terminología que nos ayudará a estudiar diferenciación de funciones de varias variables en la sección 2.3. Este material está centrado en los conceptos de conjunto abierto, límite y continuidad; los conjuntos abiertos son necesarios para entender límites, y, a su vez, los límites son necesarios para entender continuidad y diferenciabilidad.

Como sucede en cálculo elemental, no es necesario dominar completamente el concepto de límite para poder trabajar con problemas de diferenciación. Por esta razón los profesores pueden tratar el siguiente material con distintos grados de rigor. Los estudiantes deberán consultar a sus maestros acerca de la profundidad requerida.

Comenzamos la formulación del concepto de conjunto abierto mediante la definición de disco abierto. Sea $\mathbf{x}_0 \in \mathbf{R}^n$ y sea r un número real positivo. El *disco abierto* (o *bola abierta*) de radio r y centro $\mathbf{x}_0$ se define como el conjunto de todos los puntos $\mathbf{x}$ tales que $\|\mathbf{x} - \mathbf{x}_0\| < r$. Este conjunto se denota por $D_r(\mathbf{x}_0)$, y es el conjunto de puntos $\mathbf{x}$ en $\mathbf{R}^n$ cuya distancia (ver la sección 1.5) a $\mathbf{x}_0$ es menor que r. Nótese que incluimos sólo aquellas $\mathbf{x}$ para las cuales se cumple la desigualdad *estricta*. El disco $D_r(\mathbf{x}_0)$ se ilustra en la figura 2.2.1 para $n = 1$, 2 y 3. En el caso $n = 1$ y $x_0 \in \mathbf{R}$, el disco abierto $D_r(x_0)$ es el intervalo abierto $(x_0 - r, x_0 + r)$, que consta de los números $x \in \mathbf{R}$ que están *estrictamente* entre $x_0 - r$ y $x_0 + r$. En el caso $n = 2$, $\mathbf{x}_0 \in \mathbf{R}^2$, $D_r(\mathbf{x}_0)$ es el "interior" del disco de radio r con centro en $\mathbf{x}_0$. En el caso $n = 3$, $\mathbf{x}_0 \in \mathbf{R}^3$, $D_r(\mathbf{x}_0)$ es el "interior" de la bola de radio r con centro en $\mathbf{x}_0$.

DEFINICIÓN *Sea $U \subset \mathbf{R}^n$ (esto es, sea U un subconjunto de $\mathbf{R}^n$). Decimos que U es un* **conjunto abierto** *cuando para cualquier punto $\mathbf{x}_0$ en U existe algún $r > 0$ tal que $D_r(\mathbf{x}_0)$ está contenido en U; en símbolos, $D_r(\mathbf{x}_0) \subset U$ (ver la figura 2.2.2).*

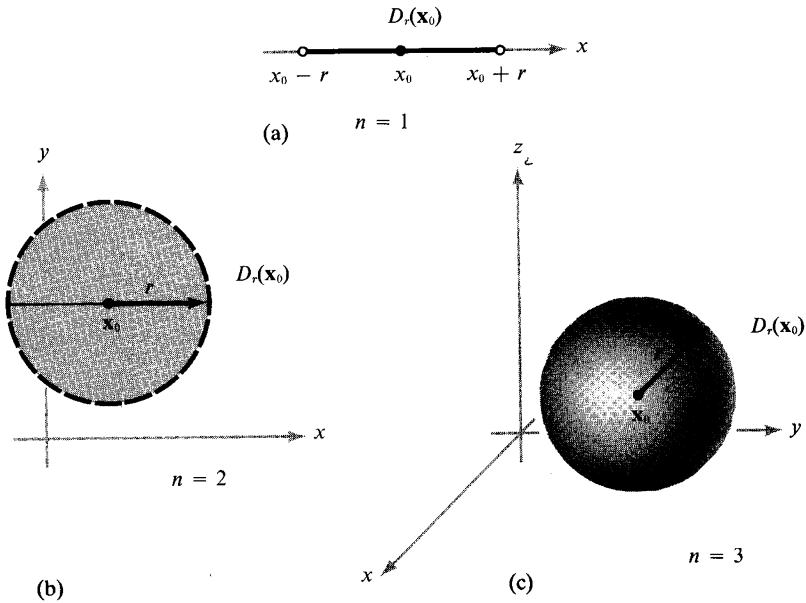


Figura 2.2.1 Apariencia de los discos $D_r(\mathbf{x}_0)$ en (a) 1, (b) 2, y (c) 3 dimensiones.

Nótese que el número $r > 0$ depende del punto $\mathbf{x}_0$, en general r disminuirá conforme $\mathbf{x}_0$ se acerca al "borde" de U. Hablando intuitivamente, un conjunto U es abierto cuando los puntos "frontera" de U no pertenecen a U. En la figura 2.2.2, la línea punteada *no* está incluida en U, y en la figura 2.2.1(c), la frontera de la esfera *no* está incluida.

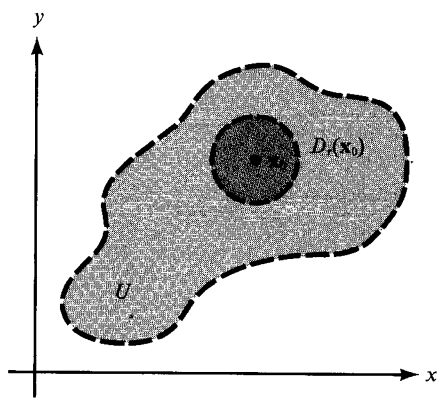


Figura 2.2.2 Un conjunto abierto U es aquel que incluye completamente algún disco $D_r(\mathbf{x}_0)$ alrededor de cada uno de sus puntos $\mathbf{x}_0$.

Además establecemos la convención de que el conjunto vacío $\emptyset$ (el conjunto que no tiene elementos) es abierto.

Hemos definido un disco abierto y un conjunto abierto. Según hemos escogido los términos, parece que un disco abierto también deberá ser un conjunto abierto. Un momento de reflexión nos hará ver que este hecho requiere una breve demostración.

TEOREMA 1 *Para cada* $\mathbf{x}_0 \in \mathbf{R}^n$ *y* $r > 0$, $D_r(\mathbf{x}_0)$ *es un conjunto abierto.*

DEMOSTRACIÓN Sea $\mathbf{x} \in D_r(\mathbf{x}_0)$, esto es, sea $\|\mathbf{x} - \mathbf{x}_0\| < r$. De acuerdo con la definición de conjunto abierto, debemos hallar un $s > 0$ tal que $D_s(\mathbf{x}) \subset D_r(\mathbf{x}_0)$. Si nos referimos a la figura 2.2.3, vemos que $s = r - \|\mathbf{x} - \mathbf{x}_0\|$ es una selección razonable; nótese que $s > 0$, pero que s se hace más chico si $\mathbf{x}$ está cerca del borde de $D_r(\mathbf{x}_0)$.

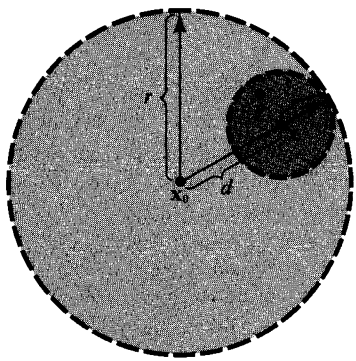


Figura 2.2.3 Geometría de la demostración de que un disco abierto es un conjunto abierto.

Para probar que $D_s(\mathbf{x}) \subset D_r(\mathbf{x}_0)$, sea $\mathbf{y} \in D_s(\mathbf{x})$; esto es, sea $\|\mathbf{y} - \mathbf{x}\| < s$. Queremos probar que también $\mathbf{y} \in D_r(\mathbf{x}_0)$. Probar esto en vista de la definición de un r-disco, equivale a demostrar que $\|\mathbf{y} - \mathbf{x}_0\| < r$. Esto se hace mediante la desigualdad del triángulo para vectores en $\mathbf{R}^n$ (ver el corolario al teorema 3, sección 1.5):

$$\begin{aligned}\|\mathbf{y} - \mathbf{x}_0\| &= \|(\mathbf{y} - \mathbf{x}) + (\mathbf{x} - \mathbf{x}_0)\| \\ &\leq \|\mathbf{y} - \mathbf{x}\| + \|\mathbf{x} - \mathbf{x}_0\| < s + \|\mathbf{x} - \mathbf{x}_0\| = r.\end{aligned}$$

De aquí que $\|\mathbf{y} - \mathbf{x}_0\| < r$. ■

El ejemplo siguiente ilustra algunas técnicas útiles para verificar que un conjunto es abierto.

EJEMPLO 1 *Probar que* $A = \{(x, y) \in \mathbf{R}^2 | x > 0\}$ *es un conjunto abierto.*

SOLUCIÓN Se muestra el conjunto en la figura 2.2.4. Intuitivamente, el conjunto es abierto pues ninguno de los puntos "frontera" $x = 0$, está contenido en el

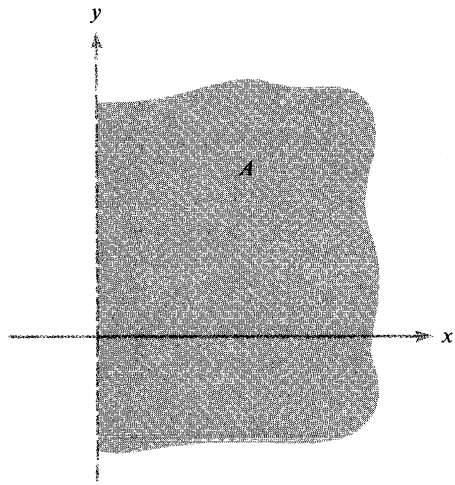


Figura 2.2.4 Problema: mostrar que A es un conjunto abierto.

conjunto. Este argumento será suficiente cuando nos hayamos acostumbrado a las ideas. Sin embargo, al principio debemos dar todos los detalles. Para probar que A es abierto, mostramos que para todo punto $(x, y) \in A$ existe un $r > 0$ tal que $D_r(x, y) \subset A$. Si $(x, y) \in A$, entonces $x > 0$. Escoger $r = x$. Si $(x_1, y_1) \in D_r(x, y)$, tenemos

$$|x_1 - x| = \sqrt{(x_1 - x)^2} \leq \sqrt{(x_1 - x)^2 + (y_1 - y)^2} < r = x,$$

y así, $x_1 - x < x$ y $x - x_1 < x$. La desigualdad anterior implica que $x_1 > 0$, esto es, $(x_1, y_1) \in A$. Entonces $D_r(x, y) \subset A$, y por lo tanto A es abierto (ver la figura 2.2.5). ▲

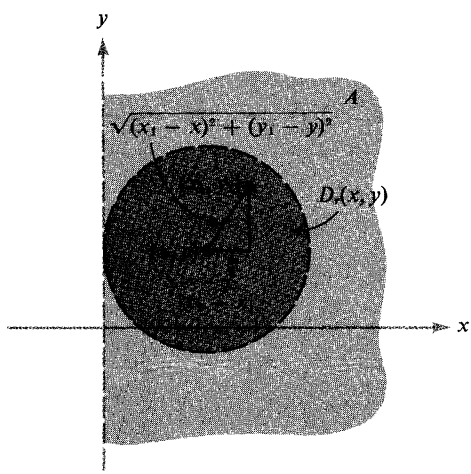


Figura 2.2.5 Construcción de un disco alrededor de un punto en A, que está completamente contenido en A.

Es útil tener un nombre especial para un conjunto abierto que contenga un punto dado $\mathbf{x}$, ya que esta idea se presenta frecuentemente al estudiar límites y continuidad. Así, designaremos como una *vecindad* de $\mathbf{x} \in \mathbf{R}^n$ a un conjunto abierto U que contiene al punto $\mathbf{x}$. Por ejemplo, $D_r(\mathbf{x}_0)$ es una vecindad de $\mathbf{x}_0$ para cualquier $r > 0$. El conjunto A en el ejemplo 1 es una vecindad del punto $\mathbf{x}_0 = (3, -10)$.

Introduzcamos formalmente el concepto de punto frontera aludido en el ejemplo 1.

DEFINICIÓN *Sea $A \subset \mathbf{R}^n$. Un punto $\mathbf{x} \in \mathbf{R}^n$ es* **punto frontera** *de A si toda vecindad de $\mathbf{x}$ contiene al menos un punto en A y al menos un punto fuera de A.*

En esta definición, $\mathbf{x}$ puede estar o no en A; si $\mathbf{x} \in A$, entonces $\mathbf{x}$ es un punto frontera si toda vecindad de $\mathbf{x}$ contiene al menos un punto que *no* esté en A (ya contiene un punto en A, a saber, $\mathbf{x}$). De manera análoga, si $\mathbf{x}$ no está en A, es un punto frontera si toda vecindad de $\mathbf{x}$ contiene al menos un punto de A.

Estaremos particularmente interesados en puntos frontera de conjuntos abiertos. Por la definición de conjunto abierto, ningún punto de un conjunto abierto A puede ser un punto frontera de A. Así, *un punto $\mathbf{x}$ es un punto frontera de un conjunto abierto A si y sólo si $\mathbf{x}$ no está en A y toda vecindad de $\mathbf{x}$ tiene intersección no vacía con A.*

Esto expresa en términos precisos la idea intuitiva de que un punto frontera de A es un punto justo en el "borde" de A. En la mayoría de los ejemplos resulta muy claro cuáles son los puntos frontera.

EJEMPLO 2 (a) Sea $A = (a, b)$ en $\mathbf{R}$. Entonces los puntos frontera de A son los puntos a y b. Un examen de la figura 2.2.6 y de la definición, lo aclaran. (Se pedirá al lector probar esto en el ejercicio 2(c).)

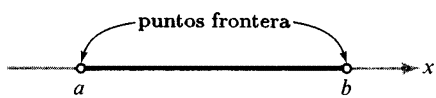


Figura 2.2.6 Los puntos frontera del intervalo (a, b).

(b) Sea $A = D_r(x_0, y_0)$ un r-disco alrededor de (x_0, y_0) en el plano. La frontera está formada por los puntos (x, y) con $(x - x_0)^2 + (y - y_0)^2 = r^2$ (figura 2.2.7).

(c) Sea $A = \{(x, y) \in \mathbf{R}^2 | x > 0\}$. Entonces la frontera de A está formada por todos los puntos sobre el eje y (ver la figura 2.2.8).

(d) Sea A el conjunto $D_r(\mathbf{x}_0)$ menos el punto x_0 (un disco "agujereado" alrededor de $\mathbf{x}_0$). Entonces $\mathbf{x}_0$ es un punto frontera de A. ▲

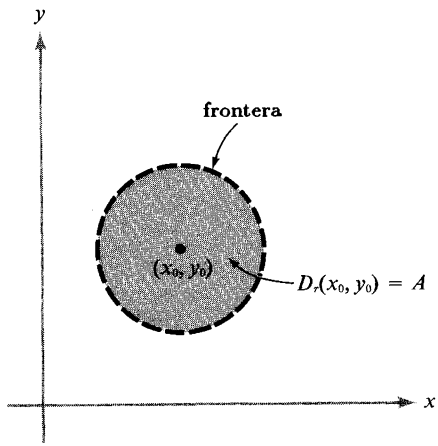


Figura 2.2.7 La frontera de A está formada por los puntos en el borde de A.

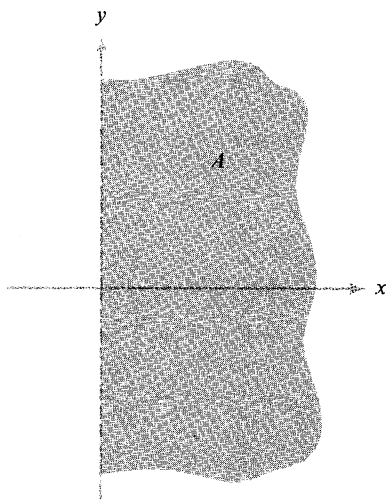


Figura 2.2.8 La frontera de A está formada por todos los puntos del eje y.

Ahora nos ocuparemos del concepto de límite. Durante toda la exposición siguiente, el dominio de definición de la función f será un conjunto abierto A. Estamos interesados en hallar el límite de f cuando $\mathbf{x} \in A$ tienda a un punto de A o a un punto frontera de A.

El lector deberá comprender el hecho de que el concepto de límite es una herramienta básica y útil para el análisis de funciones; nos permite estudiar derivadas y por lo tanto, máximos y mínimos, asíntotas, integrales impropias y otras características importantes de las funciones; también es útil para series infinitas y

sucesiones. Presentaremos una teoría de límites de funciones de varias variables que abarque, como caso particular, a la teoría de funciones de una variable.

El estudiante probablemente ya aprendió, del cálculo de una variable, una definición de $\lim_{x \to x_0} f(x) = l$ para una función $f: A \subset \mathbf{R} \to \mathbf{R}$ de un subconjunto A de números reales a los números reales. Intuitivamente, esto significa que conforme x se acerca más y más a x_0, los valores $f(x)$ se acercan más y más a l. Para colocar esta idea intuitiva sobre una base firme, precisa y manejable desde el punto de vista matemático, por lo general se introduce el "método de épsilon (ε) y delta (δ)" o el "método de las vecindades". Lo mismo sucede con las funciones de varias variables. A continuación desarrollaremos el enfoque de las vecindades para el concepto de límite. El enfoque épsilon-delta se deja como material de estudio opcional al final de esta sección.

DEFINICIÓN DE LÍMITE *Sea $f: A \subset \mathbf{R}^n \to \mathbf{R}^m$, donde A es un conjunto abierto. Sea $\mathbf{x}_0$ un punto en A o en la frontera de A, y sea V una vecindad de $\mathbf{b} \in \mathbf{R}^m$. Decimos que f está* **finalmente*** *en V* **conforme** $\mathbf{x}$ **tiende a** $\mathbf{x}_0$ *si existe una vecindad U de $\mathbf{x}_0$ tal que $\mathbf{x} \neq \mathbf{x}_0$, $\mathbf{x} \in U$ y $\mathbf{x} \in A$ implica $f(\mathbf{x}) \in V$. (En la figura 2.2.9 se ilustra el significado geométrico de esta afirmación; nótese que $\mathbf{x}_0$ no necesariamente debe estar en el conjunto A, de modo que no necesariamente está definida $f(\mathbf{x}_0)$.) Decimos que $f(\mathbf{x})$* **tiende a** $\mathbf{b}$ *cuando $\mathbf{x}$ tiende a $\mathbf{x}_0$, o, en símbolos,*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \mathbf{b} \qquad o \qquad f(\mathbf{x}) \to \mathbf{b} \qquad cuando \qquad \mathbf{x} \to \mathbf{x}_0,$$

cuando, dada **cualquier** *vecindad V de $\mathbf{b}$, f está finalmente en V conforme $\mathbf{x}$ tiende a $\mathbf{x}_0$ (esto es, "$f(\mathbf{x})$ está cerca de $\mathbf{b}$ si $\mathbf{x}$ está cerca de $\mathbf{x}_0$"). Puede ser que cuando $\mathbf{x}$ tienda a $\mathbf{x}_0$ los valores de $f(\mathbf{x})$ no se acerquen a un número particular. En este caso decimos que* $\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x})$ **no existe.**

De ahora en adelante, cuando consideremos el concepto de $\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x})$, supondremos que $\mathbf{x}_0$ pertenece a cierto conjunto abierto donde está definida f, o bien está en la frontera de dicho conjunto.

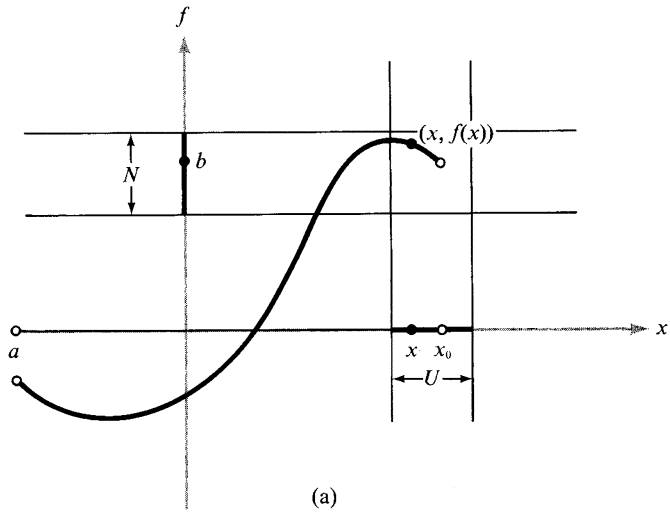
Una de las razones por las que insistimos en la definición de límite, en que $\mathbf{x} \neq \mathbf{x}_0$, será clara si recordamos, del cálculo de una variable, que nuestra intención es definir la derivada $f'(x_0)$ de una función f en un punto x_0 mediante

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

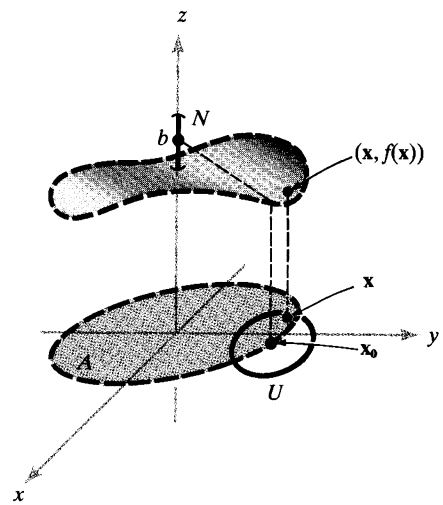
y esta expresión no está definida en $x = x_0$.

---

*Usamos "***finalmente***" como traducción de "***eventually***". El significado en castellano de la traducción literal "eventualmente" es: *incierta o casualmente* (cf. *Diccionario de la Lengua Española*. Madrid: Real Academia Española, 1984); mientras que el significado en inglés de "eventually" (y, por supuesto, con el sentido que se usa en la definición) es: *como resultado final* (cf. *The International Webster Dictionary of the English Language*. Nueva York: Tabor House, 1973). [N. del T.]



(a)



(b)

Figura 2.2.9 Límites en términos de vecindades; si $\mathbf{x}$ está en U, entonces $f(\mathbf{x})$ estará en V. (Los pequeños círculos huecos denotan el hecho de que $(a, f(a))$ y $(x_0, f(x_0))$ no están en la gráfica.)
(a) $f: A = (a, x_0) \to \mathbf{R}$. (b) $f: A = \{(x, y) | x^2 + y^2 < 1\} \to \mathbf{R}$. (La línea punteada no está en la gráfica de f.)

EJEMPLO 3 (a) Este ejemplo ilustra un límite que no existe. Considerar la función $f: \mathbf{R} \to \mathbf{R}$ definida por

$$f(x) = \begin{cases} 1 & \text{si} \quad x > 0 \\ -1 & \text{si} \quad x \leq 0 \end{cases}$$

No existe el límite $\lim_{x \to 0} f(x)$ pues hay puntos x_1 arbitrariamente cerca de 0 con $f(x_1) = 1$, y también hay puntos x_2 arbitrariamente cerca de 0 con $f(x_2) = -1$; esto es, no hay un solo número cerca del cual esté f cuando x esté cerca de 0 (ver la figura 2.2.10). Si se restringe f al dominio $(0, 1)$ o $(-1, 0)$, entonces el límite existe. ¿Pueden ver por qué?

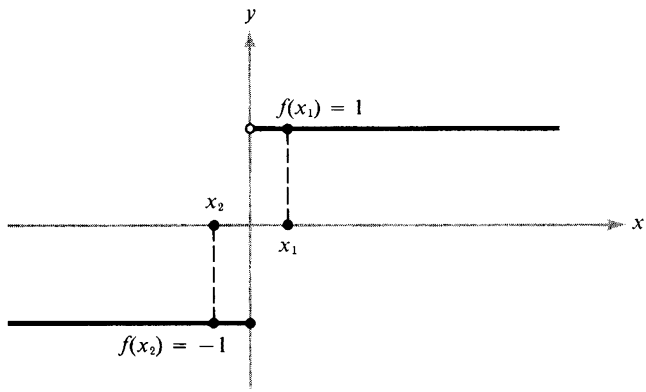


Figura 2.2.10 No existe el límite de esta función cuando $x \to 0$.

(b) Con este ejemplo se ilustra una función cuyo límite existe, pero cuyo valor límite no es igual a su valor en el punto donde se toma el límite. Definir $f: \mathbf{R} \to \mathbf{R}$ mediante

$$f(x) = \begin{cases} 0 & \text{si} \quad x \neq 0 \\ 1 & \text{si} \quad x = 0 \end{cases}$$

Es cierto que $\lim_{x \to 0} f(x) = 0$, pues para cualquier vecindad U de 0, $x \in U$ y $x \neq 0$ implica que $f(x) = 0$. Si no hubiésemos insistido en que $x \neq x_0$, entonces

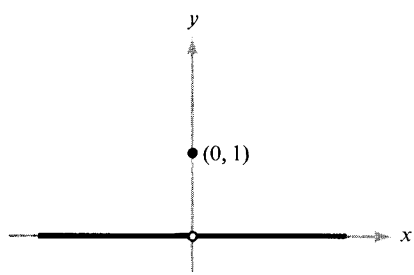


Figura 2.2.11 El límite de esta función cuando $x \to 0$ es cero.

el límite (suponiendo que usamos la definición anterior de límite sin la condición $\mathbf{x} \neq \mathbf{x}_0$) no existiría. Así, estamos realmente interesados en el valor al que se acerca f cuando $x \to 0$ (o, en general, cuando $\mathbf{x} \to \mathbf{x}_0$). En la gráfica de la figura 2.2.11 vemos que f tiende a 0 cuando $x \to 0$; no nos importa si f toma otro valor en 0, o no está definida ahí. ▲

EJEMPLO 4 *Usar la definición para verificar que se cumple el "límite obvio"* $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{x} = \mathbf{x}_0$, *donde* $\mathbf{x}$ *y* $\mathbf{x}_0 \in \mathbf{R}^n$.

SOLUCIÓN Sea f la función definida por $f(\mathbf{x}) = \mathbf{x}$, y sea V cualquier vecindad de $\mathbf{x}_0$. Debemos mostrar que $f(\mathbf{x})$ está finalmente en V cuando $\mathbf{x} \to \mathbf{x}_0$. Así, por la definición, debemos hallar una vecindad U de $\mathbf{x}_0$ con la propiedad de que si $\mathbf{x} \neq \mathbf{x}_0$ y $\mathbf{x} \in U$, entonces $f(\mathbf{x}) \in V$. Escoger $U = V$. Si $\mathbf{x} \in U$, entonces $\mathbf{x} \in V$; como $\mathbf{x} = f(\mathbf{x})$, se sigue que $f(\mathbf{x}) \in V$. Así, hemos mostrado que $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{x} = \mathbf{x}_0$. De manera análoga tenemos

$$\lim_{(x,y)\to(x_0,y_0)} \mathbf{x} = \mathbf{x}_0, \qquad \text{etc.} \quad ▲$$

A partir de ahora el estudiante puede dar por válidos, sin demostración, los límites que aprendió en cálculo de una variable. Por ejemplo, se puede usar $\lim_{x\to 1} \sqrt{x} = \sqrt{1} = 1$ y $\lim_{\theta\to 0} \operatorname{sen}\theta = \operatorname{sen} 0 = 0$.

EJEMPLO 5 (En este ejemplo se muestra otro caso, en el cual no es posible simplemente "omitir" el límite.) *Hallar* $\lim_{x\to 1} g(x)$ *donde*

$$g: x \mapsto \frac{x-1}{\sqrt{x}-1}.$$

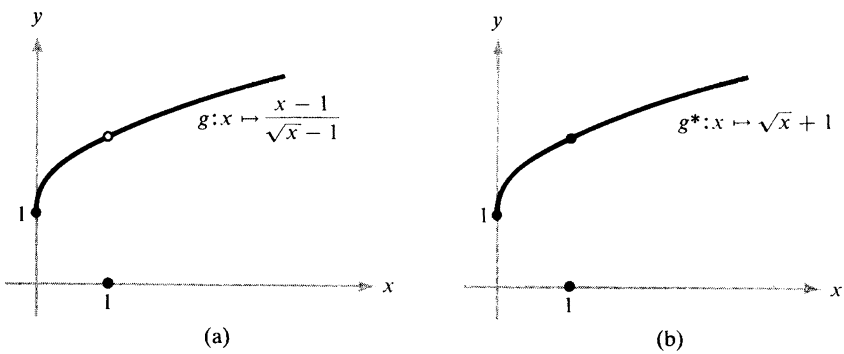


Figura 2.2.12 Estas gráficas son iguales, excepto que en la parte (a) g no está definida en $x = 1$, mientras que en la parte (b) g^* está definida para todo $x \geq 0$.

SOLUCIÓN La gráfica de esta función está en la figura 2.2.12. Vemos que $g(1)$ no está definido, pues la división entre cero no está definida. Sin embargo, si multiplicamos el numerador y el denominador de $g(x)$ por $\sqrt{x}+1$, hallamos que para todo x en el dominio de g, tenemos

$$g(x) = \frac{x-1}{\sqrt{x}-1} = \sqrt{x}+1, \qquad x \neq 1.$$

La expresión $g^*(x) = \sqrt{x}+1$ está definida y toma el valor 2 en $x = 1$; por el cálculo en una variable, $g^*(x) \to 2$ cuando $x \to 1$. Pero como $g^*(x) = g(x)$ para todo $x \geq 0$, $x \neq 1$, debemos tener también que $g(x) \to 2$ cuando $x \to 1$. ▲

Para hablar con propiedad acerca de *el* límite, debemos mostrar que f puede tener *a lo más un* límite cuando $\mathbf{x} \to \mathbf{x}_0$. Esto resulta intuitivamente claro y ahora lo enunciamos formalmente. (Ver la sección 2.7 para la demostración.)

TEOREMA 2: UNICIDAD DE LOS LÍMITES
Si $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}_1$ *y* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}_2$, *entonces* $\mathbf{b}_1 = \mathbf{b}_2$.

Para realizar cálculos prácticos con límites necesitamos algunas reglas, por ejemplo, que el límite de una suma es la suma de los límites. Estas reglas se resumen en el siguiente teorema (ver la sección 2.7 para la demostración).

TEOREMA 3 *Sean* $f: A \subset \mathbf{R}^n \to \mathbf{R}^m$, $g: A \subset \mathbf{R}^n \to \mathbf{R}^m$, $\mathbf{x}_0$ *un elemento de* A *o un punto frontera de* A, $\mathbf{b} \in \mathbf{R}^m$ *y* $c \in \mathbf{R}$; *entonces*

(i) *Si* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}$, *entonces* $\lim_{\mathbf{x}\to\mathbf{x}_0} cf(\mathbf{x}) = c\mathbf{b}$, *donde* $cf: A \to \mathbf{R}^m$ *está definida por* $\mathbf{x} \mapsto c(f(\mathbf{x}))$.

(ii) *Si* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}_1$ *y* $\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = \mathbf{b}_2$, *entonces* $\lim_{\mathbf{x}\to\mathbf{x}_0} (f+g)(\mathbf{x}) = \mathbf{b}_1+\mathbf{b}_2$, *donde* $(f+g): A \to \mathbf{R}^m$ *está definida por* $\mathbf{x} \mapsto f(\mathbf{x}) + g(\mathbf{x})$.

(iii) *Si* $m = 1$, $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = b_1$ *y* $\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = b_2$, *entonces* $\lim_{\mathbf{x}\to\mathbf{x}_0} (fg)(\mathbf{x}) = b_1 b_2$, *donde* $(fg): A \to \mathbf{R}$ *está definida por* $\mathbf{x} \mapsto f(\mathbf{x})g(\mathbf{x})$.

(iv) *Si* $m = 1$, $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = b \neq 0$ *y* $f(\mathbf{x}) \neq 0$ *para todo* $\mathbf{x} \in A$, *entonces* $\lim_{\mathbf{x}\to\mathbf{x}_0} 1/f(\mathbf{x}) = 1/b$, *donde* $1/f: A \to \mathbf{R}$ *está definida por* $\mathbf{x} \mapsto 1/f(\mathbf{x})$.

(v) *Si* $f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$ *donde* $f_i: A \to \mathbf{R}$, $i = 1, \ldots, m$, *son las funciones componentes de* f, *entonces* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b} = (b_1, \ldots, b_m)$ *si y sólo si* $\lim_{\mathbf{x}\to\mathbf{x}_0} f_i(\mathbf{x}) = b_i$ *para cada* $i = 1, \ldots, m$.

Estos resultados son intuitivamente claros. Por ejemplo, la regla (ii) no dice más que si $f(\mathbf{x})$ está cerca de $\mathbf{b}_1$ y $g(\mathbf{x})$ está cerca de $\mathbf{b}_2$ cuando $\mathbf{x}$ está cerca

de $\mathbf{x}_0$, entonces $f(\mathbf{x}) + g(\mathbf{x})$ está cerca de $\mathbf{b}_1 + \mathbf{b}_2$ cuando $\mathbf{x}$ está cerca de $\mathbf{x}_0$. El siguiente ejemplo ilustrará la situación.

EJEMPLO 6 *Sea* $f: \mathbf{R}^2 \to \mathbf{R}$, $(x, y) \mapsto x^2 + y^2 + 2$. *Calcular* $\lim_{(x,y)\to(0,1)} f(x, y)$.

SOLUCIÓN Aquí f es la suma de las tres funciones $(x, y) \mapsto x^2$, $(x, y) \mapsto y^2$, y $(x, y) \mapsto 2$. El límite de una suma es la suma de los límites, y el límite de un producto es el producto de los límites (teorema 3). Por lo tanto, usando el hecho de que $\lim_{(x,y)\to(x_0,y_0)} x = x_0$ (ejemplo 4), obtenemos

$$\lim_{(x,y)\to(x_0,y_0)} x^2 = \left(\lim_{(x,y)\to(x_0,y_0)} x\right)\left(\lim_{(x,y)\to(x_0,y_0)} x\right) = x_0^2$$

y, usando el mismo razonamiento, $\lim_{(x,y)\to(x_0,y_0)} y^2 = y_0^2$. En consecuencia,

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = 0^2 + 1^2 + 2 = 3. \qquad \blacktriangle$$

En el curso de cálculo de una variable aprendimos que el concepto de función continua está basado en la idea intuitiva de una función cuya gráfica es una curva sin romper, esto es, una curva sin *saltos*, o el tipo de curva que trazaría una partícula en movimiento o al mover la punta de un lápiz sin separarla del papel.

Para efectuar un análisis detallado de las funciones, necesitamos conceptos más precisos que estas vagas ideas. Aclaremos mediante un ejemplo. Consideremos la función $f: \mathbf{R} \to \mathbf{R}$ definida por $f(x) = -1$ si $x \leq 0$ y $f(x) = 1$ si $x > 0$. La gráfica de f se muestra en la figura 2.2.13. (El pequeño círculo hueco significa que el punto $(0, 1)$ *no* está en la gráfica de f.) Claramente, la gráfica de f está rota en $x = 0$. Considerar ahora la función $g: x \mapsto x^2$. Se muestra esta función

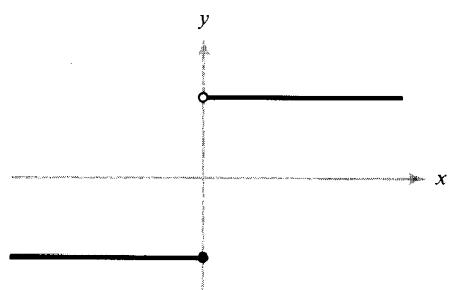


Figura 2.2.13 Esta función no es continua, pues su valor brinca súbitamente cuando x pasa por 0.

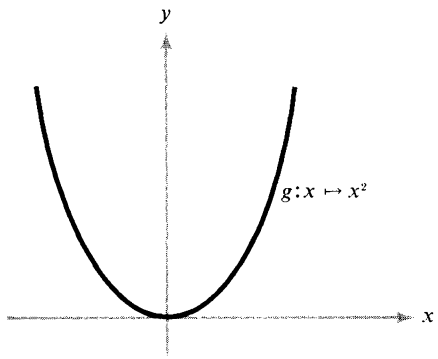


Figura 2.2.14 Esta función es continua.

en la figura 2.2.14. La gráfica de g no está rota en ningún punto. Si se examinan ejemplos de funciones como f, cuyas gráficas estén rotas en algún punto x_0, y funciones como g, cuyas gráficas no estén rotas, se ve que la diferencia principal entre ellas es que para una función como g, los valores de $g(x)$ se acercan más y más a $g(x_0)$ conforme x se acerca más y más a x_0. La misma idea sirve para funciones de varias variables. Pero la idea de más y más cerca no basta como definición matemática; así, formularemos estos conceptos de manera precisa en términos de límites.

DEFINICIÓN *Sea* $f: A \subset \mathbf{R}^n \to \mathbf{R}^m$ *una función dada con dominio* A. *Sea* $\mathbf{x}_0 \in A$. *Decimos que* f *es* ***continua en*** $\mathbf{x}_0$ *si y sólo si*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

Si decimos simplemente que f *es* ***continua****, queremos decir que* f *es continua en cada punto* $\mathbf{x}_0$ *de* A.

Como la condición $\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$ significa que $f(\mathbf{x})$ está cerca de $f(\mathbf{x}_0)$ cuando $\mathbf{x}$ está cerca de $\mathbf{x}_0$, vemos que nuestra definición corresponde, en efecto, al requerimiento de que la gráfica de f no esté rota (ver la figura 2.2.15, donde se ilustra el caso $f: \mathbf{R} \to \mathbf{R}$). El caso de varias variables es más fácil de visualizar si trabajamos con funciones con valores reales, digamos $f: \mathbf{R}^2 \to \mathbf{R}$. En este caso podemos visualizar f trazando su gráfica, formada por todos los puntos (x, y, z) en $\mathbf{R}^3$ con $z = f(x, y)$. La continuidad de f significa entonces que su gráfica no tiene "brincos" (ver la figura 2.2.16).

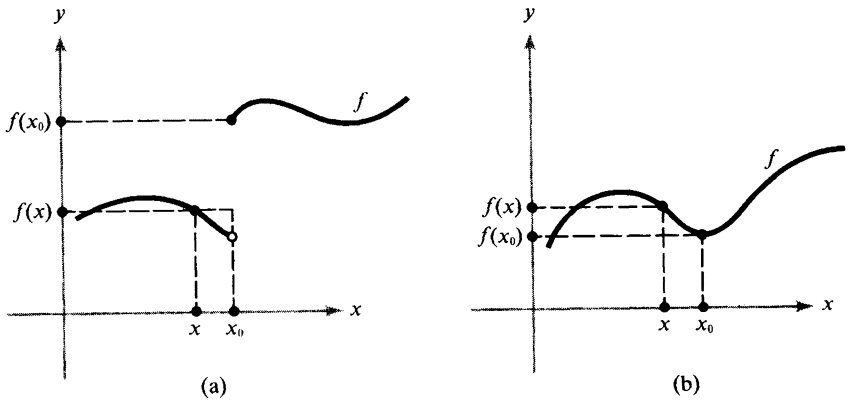


Figura 2.2.15 (a) Función discontinua para la cual no existe límite $_{x\to x_0} f(x)$; (b) función continua para la cual existe el límite y es igual a $f(x_0)$.

EJEMPLO 7 Cualquier polinomio $p(x) = a_0 + a_1x + \cdots + a_nx^n$ es continuo de **R** a **R**. En efecto, por el teorema 3 y el ejemplo 4,

$$\lim_{x\to x_0}(a_0 + a_1x + \cdots + a_nx^n) = \lim_{x\to x_0} a_0 + \lim_{x\to x_0} a_1x + \cdots + \lim_{x\to x_0} a_nx^n$$

$$= a_0 + a_1x_0 + \cdots + a_nx_0^n,$$

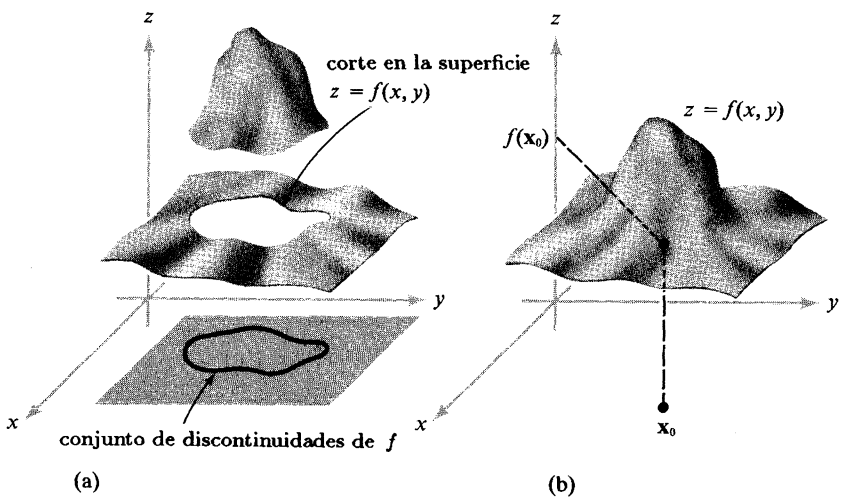


Figura 2.2.16 (a) Función discontinua de dos variables. (b) Función continua.

pues

$$\lim_{x \to x_0} x^n = \left(\lim_{x \to x_0} x\right) \cdots \left(\lim_{x \to x_0} x\right) = x_0^n. \quad \blacktriangle$$

EJEMPLO 8 Sea $f: \mathbf{R}^2 \to \mathbf{R}$, $f(x, y) = xy$. Entonces f es continua, pues, por los teoremas de límites y el ejemplo 4,

$$\lim_{(x,y) \to (x_0,y_0)} xy = \left(\lim_{(x,y) \to (x_0,y_0)} x\right)\left(\lim_{(x,y) \to (x_0,y_0)} y\right) = x_0 y_0. \quad \blacktriangle$$

Se puede ver, por el mismo método, que cualquier polinomio $p(x, y)$ en x y y es continuo.

EJEMPLO 9 La función $f: \mathbf{R}^2 \to \mathbf{R}$ definida por

$$f(x, y) = \begin{cases} 1 & \text{si} \quad x \leq 0 \quad \text{o} \quad y \leq 0 \\ 0 & \text{de no ser así} \end{cases}$$

no es continua en $(0, 0)$ o en cualquier punto sobre el eje x positivo o el eje y positivo. En efecto, si $(x_0, y_0) = \mathbf{u}$ es uno de dichos puntos y $\delta > 0$, hay puntos $(x, y) \in D_\delta(\mathbf{u})$, vecindad de $\mathbf{u}$, con $f(x, y) = 1$ y otros puntos $(x, y) \in D_\delta(\mathbf{u})$ con $f(x, y) = 0$. Así, *no* es cierto que $f(x, y) \to f(x_0, y_0) = 1$ cuando $(x, y) \to (x_0, y_0)$. $\blacktriangle$

Para probar que ciertas funciones específicas son continuas, podemos apoyarnos en los teoremas de límites (ver el teorema 3 y el ejemplo 7). Si transcribimos esos resultados en términos de continuidad, llegaremos a lo siguiente:

TEOREMA 4 *Sean $f: A \subset \mathbf{R}^n \to \mathbf{R}^m$, $g: A \subset \mathbf{R}^n \to \mathbf{R}^m$, y c un número real.*

(i) *Si f es continua en $\mathbf{x}_0$ también lo es cf, donde $(cf)(\mathbf{x}) = c[f(\mathbf{x})]$.*

(ii) *Si f y g son continuas en $\mathbf{x}_0$, también lo es $f + g$, donde $(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$.*

(iii) *Si f y g son continuas en $\mathbf{x}_0$ y $m = 1$, entonces la función producto fg definida por $(fg)(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ es continua en $\mathbf{x}_0$.*

(iv) *Si $f: A \subset \mathbf{R}^n \to \mathbf{R}$ es continua en $\mathbf{x}_0$ y no se anula en A, entonces el cociente $1/f$ es continuo en $\mathbf{x}_0$, donde $(1/f)(\mathbf{x}) = 1/f(\mathbf{x})$.**

(v) Si $f: A \subset \mathbf{R}^n \to \mathbf{R}^m$ y $f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$, entonces f es continua en $\mathbf{x}_0$ si y sólo si cada una de las funciones con valores reales $f_1, \ldots, f_m$ es continua en $\mathbf{x}_0$.

---

*Otra manera de enunciar la regla (iv) es: Si $f(\mathbf{x}_0) \neq 0$ y f es continua, entonces $f(\mathbf{x}) \neq 0$ en una vecindad de $\mathbf{x}_0$ de modo que $1/f$ está definida en esa vecindad, y $1/f$ es continua en $\mathbf{x}_0$.

EJEMPLO 10 *Sea* $f: \mathbf{R}^2 \to \mathbf{R}^2$, $(x, y) \mapsto (x^2y, (y + x^3)/(1 + x^2))$. *Mostrar que* f *es continua.*

SOLUCIÓN Para verlo es suficiente, por la propiedad (v) anterior, mostrar que cada componente es continua. Entonces, mostramos primero que $(x, y) \mapsto x^2y$ es continua. Ahora bien, $(x, y) \mapsto x$ es continua (ver el ejemplo 4), y de ahí, por (iii), $(x, y) \mapsto x^2$ es continua. Como $(x, y) \mapsto y$ es continua, por (iii), la corrrespondencia $(x, y) \mapsto x^2y$ es continua. Como $1 + x^2$ es continua y diferente de cero, por la propiedad (iv) sabemos que $1/(1+x^2)$ es continua; por lo tanto $(y+x^3)/(1+x^2)$ es un producto de funciones continuas, y por (iii) es continua. ▲

A continuación estudiaremos la *composición*, otra operación básica que puede efectuarse con funciones. Si g manda A a B y f manda B a C, la composición de g con f, o de f en g, denotada por $f \circ g$, manda A a C mediante $\mathbf{x} \mapsto f(g(\mathbf{x}))$ (ver la figura 2.2.17). Por ejemplo, $\operatorname{sen}(x^2)$ es la composición de $x \mapsto x^2$ con $y \mapsto \operatorname{sen} y$.

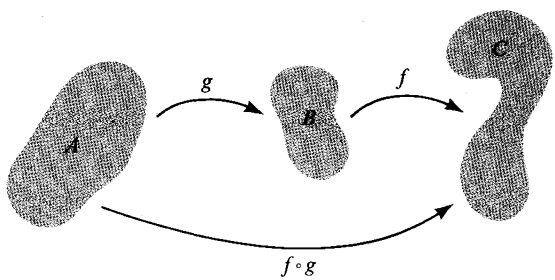


Figura 2.2.17 Composición de f en g.

TEOREMA 5 *Sean* $g: A \subset \mathbf{R}^n \to \mathbf{R}^m$ *y* $f: B \subset \mathbf{R}^m \to \mathbf{R}^p$. *Suponer que* $g(A) \subset B$, *de manera que* $f \circ g$ *está definida en* A. *Si* g *es continua en* $\mathbf{x}_0 \in A$ *y* f *es continua en* $\mathbf{y}_0 = g(\mathbf{x}_0)$, *entonces* $f \circ g$ *es continua en* $\mathbf{x}_0$.

La idea intuitiva es fácil; la demostración formal que aparece en la sección 2.7 sigue un patrón similar. Intuitivamente, debemos mostrar que conforme $\mathbf{x}$ se acerca a $\mathbf{x}_0$, $f(g(\mathbf{x}))$ se acerca a $f(g(\mathbf{x}_0))$. Pero conforme $\mathbf{x}$ se acerca a $\mathbf{x}_0$, $g(\mathbf{x})$ se acerca a $g(\mathbf{x}_0)$ (por la continuidad de g en $\mathbf{x}_0$); y conforme $g(\mathbf{x})$ se acerca a $g(\mathbf{x}_0)$, $f(g(\mathbf{x}))$ se acerca a $f(g(\mathbf{x}_0))$ (por la continuidad de f en $g(\mathbf{x}_0)$).

EJEMPLO 11 *Sea* $f(x, y, z) = (x^2+y^2+z^2)^{30}+\operatorname{sen} z^3$. *Mostrar que* f *es continua.*

SOLUCIÓN Aquí podemos escribir f como suma de las dos funciones $(x^2+y^2+z^2)^{30}$ y $\operatorname{sen} z^3$, de modo que basta mostrar que cada una es continua. La primera

es la composición de $(x, y, z) \mapsto (x^2 + y^2 + z^2)$ con $u \mapsto u^{30}$, y la segunda es la composición de $(x, y, z) \mapsto z^3$ con $u \mapsto \operatorname{sen} u$, y tenemos la continuidad por el teorema 5. ▲

Suplemento de la Sección 2.2: LÍMITES EN TÉRMINOS DE ÉPSILON Y DELTA

Ahora enunciamos un útil teorema que formula el concepto de límite en términos de épsilons y deltas y que con frecuencia se considera la *definición* de límite. Es otra manera de precisar el enunciado intuitivo de que "$f(\mathbf{x})$ se acerca a $\mathbf{b}$ cuando $\mathbf{x}$ se acerca a $\mathbf{x}_0$". El lector comprenderá mejor esta exposición si la considera a la luz de los ejemplos ya presentados.

TEOREMA 6 *Sea* $f: A \subset \mathbf{R}^n \to \mathbf{R}^m$ *y sea* $\mathbf{x}_0$ *un elemento de* A *o un punto frontera de* A. *Entonces* $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}$ *si y sólo si para todo número* $\varepsilon > 0$ *existe un* $\delta > 0$ *tal que para cualquier* $\mathbf{x} \in A$ *que satisfaga* $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$, *tenemos* $\|f(\mathbf{x}) - \mathbf{b}\| < \varepsilon$ *(ver la figura 2.2.18).*

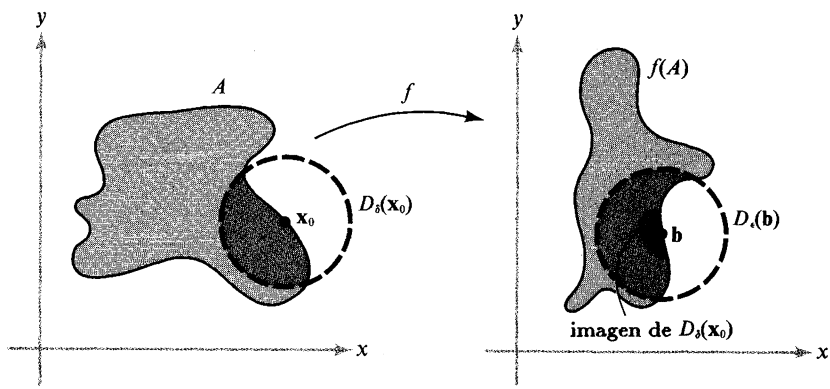


Figura 2.2.18 Geometría de la definición ε-δ de límite.

Este teorema se probará en la sección 2.7

Para ilustrar la metodología de la técnica épsilon-delta en el teorema 6, consideremos los ejemplos siguientes.

EJEMPLO 12 Mostrar que $\lim_{(x,y)\to(0,0)} x = 0$.

SOLUCIÓN Nótese que si $\delta > 0$, $\|(x, y) - (0, 0)\| = \sqrt{x^2 + y^2} < \delta$ implica que $|x-0| = |x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2} < \delta$. Así, si $\|(x, y)-(0, 0)\| < \delta$, entonces $|x-0|$ también es menor que δ. Dado $\varepsilon > 0$ se requiere que hallemos $\delta > 0$ (generalmente depende

de ε) con la propiedad de que $0 < \|(x,y) - (0,0)\| < \delta$ implique que $|x - 0| < \varepsilon$. ¿Qué δ vamos a escoger? En los cálculos anteriores vemos que si escogemos $\delta = \varepsilon$, entonces $\|(x,y) - (0,0)\| < \delta$ implica $|x - 0| < \varepsilon$. Esto muestra que $\lim_{(x,y)\to(0,0)} x = 0$. También pudimos haber escogido $\delta = \varepsilon/2$ o $\varepsilon/3$, pero basta hallar un δ para satisfacer los requerimientos de la definición de límite. ▲

EJEMPLO 13 *Considerar la función*

$$f(x,y) = \frac{\operatorname{sen}(x^2 + y^2)}{x^2 + y^2}.$$

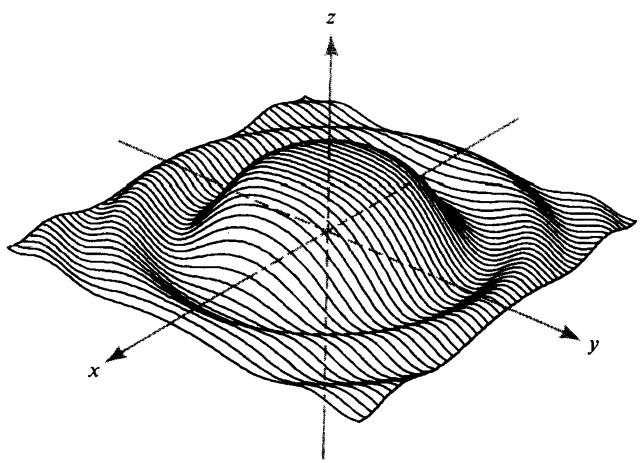
Aunque f no está definida en $(0,0)$, determinar si $f(x,y)$ tiende a algún número cuando (x,y) tiende a $(0,0)$.

SOLUCIÓN Por cálculo elemental sabemos que

$$\lim_{\alpha\to 0} \frac{\operatorname{sen}\alpha}{\alpha} = 1.$$

Por lo tanto, es razonable pensar que

$$\lim_{\mathbf{v}\to(0,0)} f(\mathbf{v}) = \lim_{\mathbf{v}\to(0,0)} \frac{\operatorname{sen}\|\mathbf{v}\|^2}{\|\mathbf{v}\|^2} = 1.$$



$$z = \frac{\operatorname{sen}(x^2 + y^2)}{(x^2 + y^2)} \qquad -3 \le x \le 3, -3 \le y \le 3$$

Figura 2.2.19 Gráfica generada por computadora.

En efecto, como $\lim_{\alpha \to 0} (\operatorname{sen} \alpha)/\alpha = 1$, dado $\varepsilon > 0$ hallamos un $\delta > 0$, con $1 > \delta > 0$, tal que $0 < |\alpha| < \delta$ implica que $|(\operatorname{sen} \alpha)/\alpha - 1| < \varepsilon$. Si $0 < \|\mathbf{v}\| < \delta$, entonces $0 < \|\mathbf{v}\|^2 < \delta^2 < \delta$, y por lo tanto

$$|f(\mathbf{v}) - 1| = \left| \frac{\operatorname{sen} \|\mathbf{v}\|^2}{\|\mathbf{v}\|^2} - 1 \right| < \varepsilon.$$

Así, $\lim_{\mathbf{v} \to (0,0)} f(\mathbf{v}) = 1$. En efecto, si graficamos en una computadora $[\operatorname{sen}(x^2+y^2)]/(x^2+y^2)$, obtenemos una gráfica bien portada cerca de $(0,0)$ (figura 2.2.19). ▲

EJEMPLO 14 *Mostrar que*

$$\lim_{(x,y) \to (0,0)} \frac{x^2}{\sqrt{x^2+y^2}} = 0.$$

SOLUCIÓN Debemos mostrar que $x^2/\sqrt{x^2+y^2}$ es pequeño cuando (x,y) está cerca del origen. Para ello usamos la siguiente desigualdad:

$$0 \leq \frac{x^2}{\sqrt{x^2+y^2}} \leq \frac{x^2+y^2}{\sqrt{x^2+y^2}} \qquad (\text{como } y^2 \geq 0)$$

$$= \sqrt{x^2+y^2}.$$

Dado $\varepsilon > 0$, escoger $\delta = \varepsilon$. Entonces $\|(x,y) - (0,0)\| = \|(x,y)\| = \sqrt{x^2+y^2}$, de modo que $\|(x,y) - (0,0)\| < \delta$ implica que

$$\left| \frac{x^2}{\sqrt{x^2+y^2}} - 0 \right| = \frac{x^2}{\sqrt{x^2+y^2}} \leq \sqrt{x^2+y^2} = \|(x,y) - (0,0)\| < \delta = \varepsilon.$$

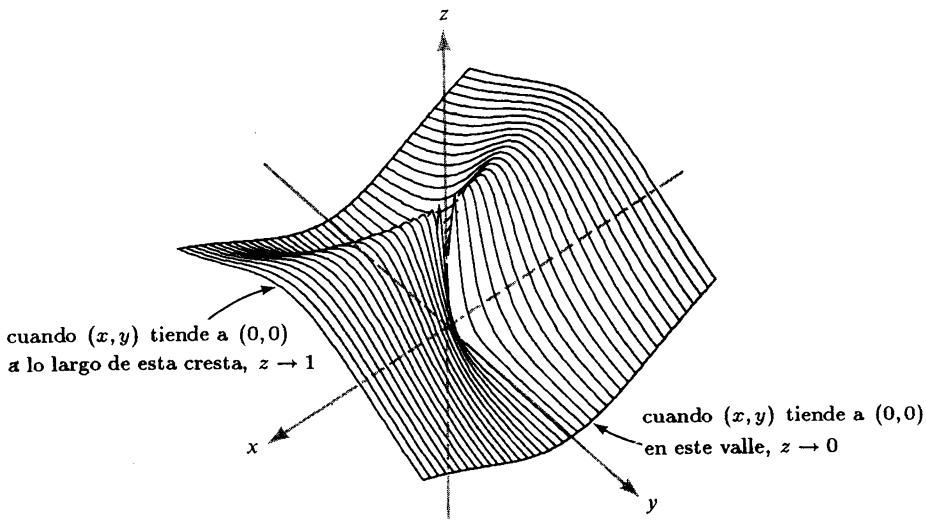
Así, se cumplen las condiciones del teorema 6 y se verifica el límite. ▲

EJEMPLO 15 *¿Existe* $\lim_{(x,y) \to (0,0)} x^2/(x^2+y^2)$*?*

SOLUCIÓN Si existe el límite, $x^2/(x^2+y^2)$ debería aproximarse a un valor definido, digamos a, cuando (x,y) se acerque a $(0,0)$. En particular, si (x,y) tiende a cero a lo largo de cualquier trayectoria, entonces $x^2/(x^2+y^2)$ debería tender al valor límite a. Si (x,y) tiende a $(0,0)$ a lo largo de la recta $y = 0$, claramente el valor límite es 1 (basta hacer $y = 0$ en la expresión anterior para obtener $x^2/x^2 = 1$). Si (x,y) tiende a $(0,0)$ a lo largo de la recta $x = 0$, el valor límite es

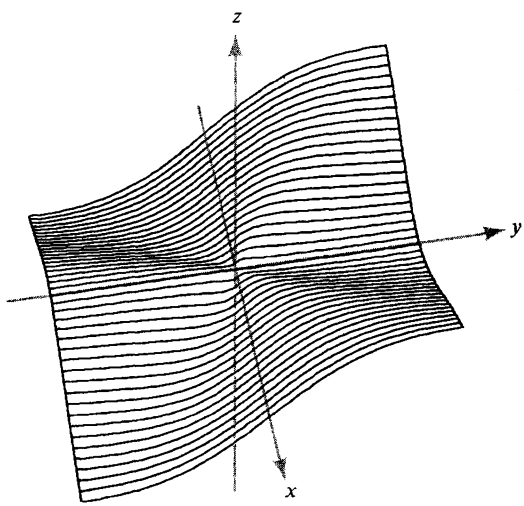
$$\lim_{(x,y) \to (0,0)} \frac{0^2}{0^2+y^2} = 0 \neq 1.$$

Por lo tanto, no existe $\lim_{(x,y) \to (0,0)} x^2/(x^2+y^2)$ (ver la figura 2.2.20). ▲



$$z = \frac{x^2}{x^2 + y^2} \qquad -1 \le x \le 1, -1 \le y \le 1$$

Figura 2.2.20 Esta función no tiene límite en $(0, 0)$.



$$z = \frac{x^2 y}{x^2 + y^2}$$

Figura 2.2.21 Esta función tiene límite 0 en $(0, 0)$.

EJEMPLO 16 *Probar que (ver la figura 2.2.21)*

$$\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2} = 0.$$

SOLUCIÓN En efecto, nótese que

$$\left|\frac{x^2y}{x^2+y^2}\right| \leq \left|\frac{x^2y}{x^2}\right| = |y|.$$

Así, dado $\varepsilon > 0$, escoger $\delta = \varepsilon$; entonces $0 < \|(x,y)-(0,0)\| = \sqrt{x^2+y^2} < \delta$ implica $|y| < \delta$, y así,

$$\left|\frac{x^2y}{x^2+y^2} - 0\right| < \varepsilon. \quad \blacktriangle$$

Usando la notación épsilon-delta llegamos a la siguiente reformulación de la definición de continuidad.

TEOREMA 7 *Sea $f: A \subset \mathbf{R}^n \to \mathbf{R}^m$ una función dada. Entonces f es continua en $\mathbf{x}_0 \in A$ si y sólo si para todo número $\varepsilon > 0$ existe un número $\delta > 0$ tal que*

$$\mathbf{x} \in A \text{ y } \|\mathbf{x}-\mathbf{x}_0\| < \delta \quad \text{implica} \quad \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \varepsilon.$$

La demostración es casi inmediata. Nótese que en el teorema 6 insistimos en que $0 < \|\mathbf{x}-\mathbf{x}_0\|$, esto es, $\mathbf{x} \neq \mathbf{x}_0$. Esto *no* se impone aquí; en efecto, ciertamente la conclusión del teorema 7 es válida cuando $\mathbf{x} = \mathbf{x}_0$, de modo que no es necesario excluir este caso. Aquí si nos interesa el valor de f en $\mathbf{x}_0$; queremos que en los puntos cercanos f esté cerca de este valor.

EJERCICIOS

En los ejercicios siguientes el lector puede suponer que las funciones exponencial, seno y coseno son continuas, y puede usar libremente las técnicas del cálculo de una variable.

1. Mostrar que los siguientes subconjuntos del plano son abiertos:

(a) $A = \{(x,y) | -1 < x < 1, -1 < y < 1\}$

(b) $B = \{(x,y) | y > 0\}$

(c) $C = \{(x,y) | 2 < x^2 + y^2 < 4\}$

(d) $A \cup B \cup C$, donde los conjuntos A, B y C se definen como en las parte (a), (b), y (c)

(e) $D = \{(x,y) | x \neq 0 \text{ y } y \neq 0\}$

2. (a) Probar que para $\mathbf{x} \in \mathbf{R}^n$ y $s < t$, $D_s(\mathbf{x}) \subset D_t(\mathbf{x})$.

(b) Probar que si U y V son vecindades de $\mathbf{x} \in \mathbf{R}^n$, entonces también lo son $U \cap V$ y $U \cup V$.

(c) Probar que los puntos frontera de un intervalo abierto $(a,b) \subset \mathbf{R}$ son los puntos a y b.

*3. Usar la formulación ε-δ de los límites para probar que $x^2 \to 4$ cuando $x \to 2$. Dar una demostración más corta usando el teorema 3.

4. Calcular los límites siguientes:
   (a) $\lim_{(x,y)\to(0,1)} x^3 y$
   (b) $\lim_{(x,y)\to(0,1)} e^x y$
   (c) $\lim_{x\to 0} \dfrac{\operatorname{sen}^2 x}{x}$
   (d) $\lim_{x\to 0} \dfrac{\operatorname{sen}^2 x}{x^2}$

5. Calcular los límites siguientes:
   (a) $\lim_{x\to 3}(x^2 - 3x + 5)$
   (b) $\lim_{x\to 0} \operatorname{sen} x$
   (c) $\lim_{h\to 0} \dfrac{(x+h)^2 - x^2}{h}$
   (d) $\lim_{h\to 0} \dfrac{e^h - 1}{h}$
   (e) $\lim_{x\to 0} \dfrac{\cos x - 1}{x^2}$

6. Calcular los límites siguientes, si es que existen:
   (a) $\lim_{(x,y)\to(0,0)} (x^2 + y^2 + 3)$
   (b) $\lim_{(x,y)\to(0,0)} \dfrac{xy}{x^2 + y^2 + 2}$
   (c) $\lim_{(x,y)\to(0,0)} \dfrac{e^{xy}}{x+1}$
   (d) $\lim_{(x,y)\to(0,0)} \dfrac{\cos x - 1 - x^2/2}{x^4 + y^4}$
   (e) $\lim_{(x,y)\to(0,0)} \dfrac{(x-y)^2}{x^2 + y^2}$

7. Calcular $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$, si es que existe, para los casos siguientes:
   (a) $f: \mathbf{R} \to \mathbf{R}, x \mapsto |x|, x_0 = 1$
   (b) $f: \mathbf{R}^n \to \mathbf{R}, \mathbf{x} \mapsto \|\mathbf{x}\|$, $\mathbf{x}_0$ arbitrario
   (c) $f: \mathbf{R} \to \mathbf{R}^2, x \mapsto (x^2, e^x), x_0 = 1$
   (d) $f: \mathbf{R}^2 \backslash \{(0,0)\} \to \mathbf{R}^2, (x,y) \mapsto (\operatorname{sen}(x-y), e^{x(y+1)} - x - 1)/\|(x,y)\|, \mathbf{x}_0 = (0,0)$

8. Sea $A \subset \mathbf{R}^2$ el disco unitario abierto $D_1(0,0)$ con el punto $\mathbf{x}_0 = (1,0)$ añadido, y sea $f: A \to \mathbf{R}$, $\mathbf{x} \mapsto f(\mathbf{x})$ la función constante $f(\mathbf{x}) = 1$. Mostrar que $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = 1$.

9. Mostrar que $\lim_{(x,y)\to(0,0)} (\operatorname{sen} xy)/xy = 1$.

10. Mostrar que la correspondencia $f: \mathbf{R} \to \mathbf{R}$, $x \mapsto x^2 e^x/(2 - \operatorname{sen} x)$ es continua.

11. Sea $f(x,y,z) = (x^2 + 3y^2)/(x+1)$. Calcular

$$\lim_{(x,y,z)\to(0,0,0)} f(x,y,z).$$

12. Mostrar que $f: \mathbf{R} \to \mathbf{R}$, $x \mapsto (1-x)^8 + \cos(1 + x^3)$ es continua.

13. Si $f: \mathbf{R}^n \to \mathbf{R}$ y $g: \mathbf{R}^n \to \mathbf{R}$ son continuas, mostrar que las funciones

$$f^2 g: \mathbf{R}^n \to \mathbf{R}, \mathbf{x} \mapsto [f(\mathbf{x})]^2 g(\mathbf{x})$$

y

$$f^2 + g: \mathbf{R}^n \to \mathbf{R}, \mathbf{x} \mapsto [f(\mathbf{x})]^2 + g(\mathbf{x})$$

son continuas.

14. Probar que $f: \mathbf{R}^2 \to \mathbf{R}$, $(x, y) \mapsto ye^x + \operatorname{sen} x + (xy)^4$ es continua.

15. (a) ¿Se puede hacer continua $[\operatorname{sen}(x+y)]/(x+y)$ definiéndola de manera adecuada en $(0,0)$?

(b) ¿Se puede hacer continua $xy/(x^2+y^2)$ definiéndola de manera adecuada en $(0,0)$?

16. (a) Usar la regla de l'Hôpital para calcular $\lim_{x \to 0} \dfrac{\operatorname{sen} 2x - 2x}{x^3}$.

(b) ¿Existe $\lim_{(x,y) \to (0,0)} \dfrac{\operatorname{sen} 2x - 2x + y}{x^3 + y}$?

17. Suponer que $\mathbf{x}$ y $\mathbf{y}$ están en $\mathbf{R}^n$ y que $\mathbf{x} \neq \mathbf{y}$. Mostrar que existe una función continua $f: \mathbf{R}^n \to \mathbf{R}$ con $f(\mathbf{x}) = 1$, $f(\mathbf{y}) = 0$, y $0 \leq f(\mathbf{z}) \leq 1$ para todo $\mathbf{z}$ en $\mathbf{R}^n$.

***18.** Sea $f: A \subset \mathbf{R}^n \to \mathbf{R}$ y sea $\mathbf{x}_0$ un punto frontera de A. Decimos que $\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \infty$ si para todo $N > 0$ existe $\delta > 0$ tal que $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$ implica $f(\mathbf{x}) > N$.

(a) Probar que $\lim_{x \to 1} (x-1)^{-2} = \infty$

(b) Probar que $\lim_{x \to 0} 1/|x| = \infty$. ¿Es cierto que $\lim_{x \to 0} 1/x = \infty$?

(c) Probar que $\lim_{(x,y) \to (0,0)} 1/(x^2+y^2) = \infty$.

***19.** Sea $f: \mathbf{R} \to \mathbf{R}$ una función. Escribimos $\lim_{x \to b^-} f(x) = L$ y decimos que L es el *límite por la izquierda* de f en b, si para todo $\varepsilon > 0$, existe $\delta > 0$ tal que $x < b$ y $0 < |x - b| < \delta$ implica $|f(x) - L| < \varepsilon$.

(a) Formular una definición de *límite por la derecha*, o $\lim_{x \to b^+} f(x)$.

(b) Hallar $\lim_{x \to 0^-} 1/(1 + e^{1/x})$ y $\lim_{x \to 0^+} 1/(1 + e^{1/x})$.

(c) Esbozar la gráfica de $1/(1 + e^{1/x})$.

***20.** Mostrar que f es continua en $\mathbf{x}_0$ si y sólo si

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \|f(\mathbf{x}) - f(\mathbf{x}_0)\| = 0.$$

*21. Sea $f: A \subset \mathbf{R}^n \to \mathbf{R}^m$ que satisface $\|f(\mathbf{x}) - f(\mathbf{y})\| \leq K\|\mathbf{x} - \mathbf{y}\|^\alpha$ para todo $\mathbf{x}$ y $\mathbf{y}$ en A para constantes positivas K y α. Mostrar que f es continua. (Dichas funciones se llaman *Hölder-continuas* o, si $\alpha = 1$, *Lipschitz-continuas*.)

*22. Mostrar que $f: \mathbf{R}^n \to \mathbf{R}^m$ es continua en todos los puntos si y sólo si la imagen inversa de todo abierto es abierta.

*23. (a) Probarque existe un número $\delta > 0$ tal que si $|a| < \delta$, entonces $|a^3 + 3a^2 + a| < 1/100$.

(b) Probar que existe un número $\delta > 0$ tal que si $x^2 + y^2 < \delta^2$, entonces

$$|x^2 + y^2 + 3xy + 180xy^5| < 1/10,000.$$

2.3 DIFERENCIACIÓN

En la sección 2.1 consideramos algunos métodos para graficar funciones. Mediante sólo estos métodos puede ser imposible calcular suficiente información para comprender incluso las características generales de una función complicada. Por el cálculo elemental sabemos que el concepto de derivada puede ser de mucha ayuda en esta tarea; por ejemplo, nos permite localizar máximos y mínimos, y calcular tasas de cambio. La derivada, además de éstas, tiene otras aplicaciones, como seguramente ya lo habrá descubierto el estudiante en su curso de cálculo elemental.

Intuitivamente ya sabemos, por nuestro trabajo con la sección 2.2, que una función continua no tiene la gráfica rota. Una función diferenciable de $\mathbf{R}^2$ a $\mathbf{R}$ debe ser tal que su gráfica no esté rota, pero además debe tener bien definido un plano tangente a la gráfica en cada punto. Así, no debe haber dobleces, esquinas

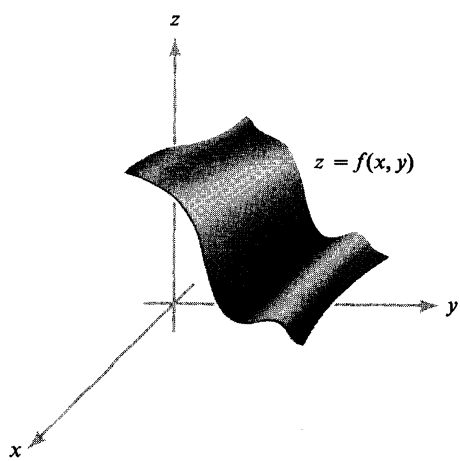


Figura 2.3.1 Gráfica suave.

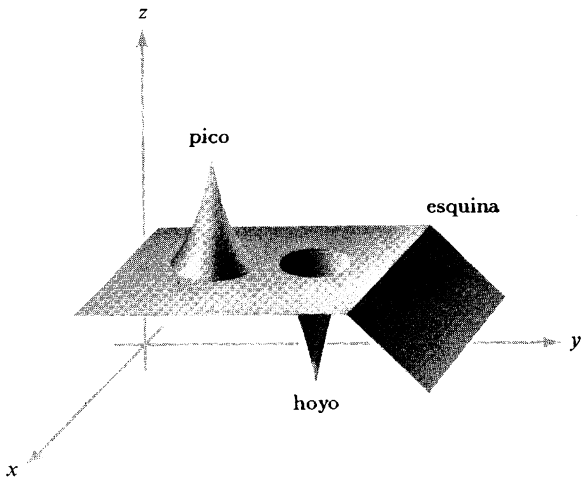


Figura 2.3.2 Esta gráfica no es suave.

o picos en la gráfica (ver las figuras 2.3.1 y 2.3.2). En otras palabras, la gráfica debe ser *suave*.

Para precisar estas ideas necesitamos una definición sensata de lo que entendemos por "$f(x_1, \ldots, x_n)$ es diferenciable en $\mathbf{x} = (x_1, \ldots, x_n)$". En realidad esta definición no es tan sencilla, como pudiera pensarse. Para avanzar en esa dirección, introduzcamos el concepto de *derivada parcial*. Este concepto se basa en nuestro conocimiento del cálculo en una variable. (En este momento se recomienda revisar rápidamente la definición de derivada en un libro de cálculo en una variable.)

DEFINICIÓN *Sean $U \subset \mathbf{R}^n$ un conjunto abierto y $f: U \subset \mathbf{R}^n \to \mathbf{R}$ una función con valores reales. Entonces $\partial f/\partial x_1, \ldots, \partial f/\partial x_n$, las* **derivadas parciales** *de f respecto a la primera, segunda, ..., n-ésima variable son las funciones con valores reales, de n variables, las cuales, en el punto $(x_1, \ldots, x_n) = \mathbf{x}$, están definidas por*

$$\frac{\partial f}{\partial x_j}(x_1, \ldots, x_n) = \lim_{h \to 0} \frac{f(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}$$

$$= \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{e}_j) - f(\mathbf{x})}{h}$$

si existen los límites, donde $1 \leq j \leq n$ y $\mathbf{e}_j$ es el j-ésimo vector de la base usual, definido por $\mathbf{e}_j = (0, \ldots, 1, \ldots, 0)$, con el 1 en el j-ésimo lugar (ver la sección 1.5).

En otras palabras, $\partial f/\partial x_j$ es simplemente la derivada de f respecto a la variable x_j, manteniendo las otras variables fijas. Si $f: \mathbf{R}^3 \to \mathbf{R}$, con frecuencia usaremos la notación $\partial f/\partial x$, $\partial f/\partial y$, y $\partial f/\partial z$ en lugar de $\partial f/\partial x_1$, $\partial f/\partial x_2$, y

$\partial f/\partial x_3$. Si $f: U \subset \mathbf{R}^n \to \mathbf{R}^m$, entonces podemos escribir

$$f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)),$$

de modo que podemos hablar de las derivadas parciales de cada componente; por ejemplo, $\partial f_m/\partial x_n$ es la derivada parcial de la m-ésima componente con respecto a x_n, la n-ésima variable.

EJEMPLO 1 *Si $f(x,y) = x^2y + y^3$, hallar $\partial f/\partial x$ y $\partial f/\partial y$.*

SOLUCIÓN Para hallar $\partial f/\partial x$ mantenemos y constante (piénsenla como si fuera un número, digamos 1) y diferenciamos sólo respecto a x; entonces

$$\frac{\partial f}{\partial x} = \frac{d(x^2y + y^3)}{dx} = 2xy.$$

De manera análoga, para hallar $\partial f/\partial y$ mantenemos x constante y diferenciamos sólo respecto a y:

$$\frac{\partial f}{\partial y} = \frac{d(x^2y + y^3)}{dy} = x^2 + 3y^2 \qquad \blacktriangle$$

Para indicar que una derivada parcial ha de evaluarse en algún punto particular, por ejemplo en (x_0, y_0), escribimos

$$\frac{\partial f}{\partial x}(x_0, y_0) \quad \text{o} \quad \left.\frac{\partial f}{\partial x}\right|_{x=x_0, y=y_0} \quad \text{o} \quad \left.\frac{\partial f}{\partial x}\right|_{(x_0, y_0)}$$

Cuando escribamos $z = f(x,y)$ para denotar la variable dependiente, con frecuencia escribiremos $\partial z/\partial x$ en lugar de $\partial f/\partial x$. Estrictamente hablando, éste es un abuso de notación, pero es una práctica común usar de manera indistinta estas dos notaciones. (Ver el ejercicio 24, en la sección 2.4 para una ilustración de los peligros que conlleva usar notación descuidada.)

EJEMPLO 2 *Si $z = \cos xy + x \cos y = f(x,y)$, hallar las derivadas parciales $(\partial z/\partial x)(x_0, y_0)$ y $(\partial z/\partial y)(x_0, y_0)$.*

SOLUCIÓN Primero fijamos y_0 y diferenciamos respecto a x, obteniendo

$$\begin{aligned}\frac{\partial z}{\partial x}(x_0, y_0) &= \left.\frac{d(\cos xy_0 + x \cos y_0)}{dx}\right|_{x=x_0} \\ &= (-y_0 \operatorname{sen} xy_0 + \cos y_0)|_{x=x_0} \\ &= -y_0 \operatorname{sen} x_0y_0 + \cos y_0.\end{aligned}$$

De manera análoga, fijamos x_0 y diferenciamos respecto a y para obtener

$$\begin{aligned}\frac{\partial z}{\partial y}(x_0, y_0) &= \left.\frac{d(\cos x_0 y + x_0 \cos y)}{dy}\right|_{y=y_0} \\ &= (-x_0 \operatorname{sen} x_0 y - x_0 \operatorname{sen} y)|_{y=y_0} \\ &= -x_0 \operatorname{sen} x_0 y_0 - x_0 \operatorname{sen} y_0. \quad \blacktriangle\end{aligned}$$

EJEMPLO 3 *Hallar* $\partial f/\partial x$ *si* $f(x,y) = xy/\sqrt{x^2+y^2}$.

SOLUCIÓN Por la regla del cociente

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{y\sqrt{x^2+y^2} - xy(x/\sqrt{x^2+y^2})}{x^2+y^2} \\ &= \frac{y(x^2+y^2) - x^2 y}{(x^2+y^2)^{3/2}} = \frac{y^3}{(x^2+y^2)^{3/2}}. \quad \blacktriangle\end{aligned}$$

Resulta insuficiente una definición de diferenciabilidad que requiera sólo de la existencia de las derivadas parciales. No se cumplirían muchos de los resultados usuales, como la regla de la cadena para varias variables, tal como lo muestra el ejemplo 4. Más adelante veremos cómo rectificar esta situación.

EJEMPLO 4 Sea $f(x,y) = x^{1/3}y^{1/3}$. Por definición,

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0,$$

y, de manera similar, $(\partial f/\partial y)(0,0) = 0$ (¡no se trata de formas indeterminadas!). Es necesario usar la definición original de derivadas parciales, pues las funciones $x^{1/3}$ y $y^{1/3}$ no son diferenciables en 0. Supongamos que restringimos f a la recta $y = x$ para obtener $f(x,x) = x^{2/3}$ (ver la figura 2.3.3). Podemos ver la substitución $y = x$ como la composición de la función $g\colon \mathbf{R} \to \mathbf{R}^2$, definida por $g(x) = (x,x)$, con $f\colon \mathbf{R}^2 \to \mathbf{R}$, definida por $f(x,y) = x^{1/3}y^{1/3}$.

Así, la composición $f \circ g$ está dada por $(f \circ g)(x) = x^{2/3}$. Cada componente de g es diferenciable en x, y f tiene derivadas parciales en $(0,0)$, pero $f \circ g$ *no es diferenciable* en $x = 0$, en el sentido del cálculo en una variable. En otras palabras, la composición de f y g no es diferenciable, en contraste con el cálculo de funciones en una variable, donde la composición de funciones diferenciables *es* diferenciable. Más adelante daremos una definición de diferenciabilidad que tiene la agradable consecuencia de que la composición de funciones diferenciables es diferenciable.

Hay otra razón para no estar satisfechos con la simple existencia de las derivadas parciales de $f(x,y) = x^{1/3}y^{1/3}$; no hay plano tangente, en un sentido razonable, a la gráfica en $(0,0)$. El plano xy es tangente a la gráfica a lo largo de los ejes x y y pues f tiene pendiente cero en $(0,0)$ a lo largo de estos ejes; esto es, $\partial f/\partial x = 0$ y $\partial f/\partial y = 0$ en $(0,0)$. Así, si hubiera plano tangente, debería ser el plano xy. Sin embargo, como resulta evidente en la figura 2.3.3, el plano xy no es tangente a la gráfica en otras direcciones, pues la gráfica tiene una severa arruga de modo que no se puede decir que el plano xy es tangente a la gráfica de f. ▲

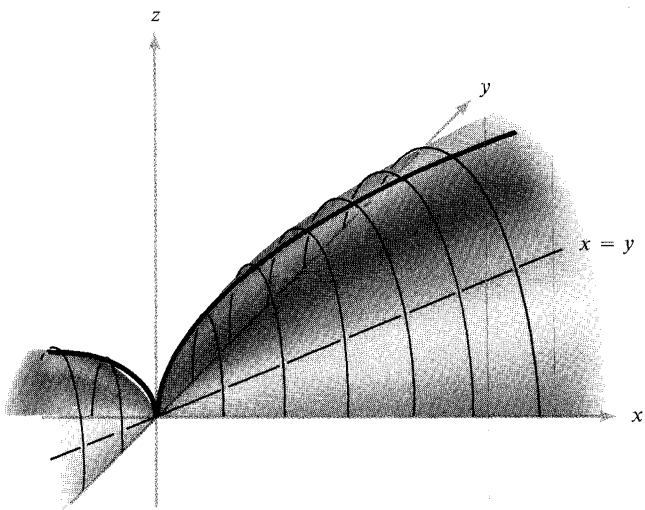


Figura 2.3.3 Parte "superior" de la gráfica de $x^{1/3}y^{1/3}$.

Para "motivar" la definición de diferenciabilidad, calculemos cómo debería ser la ecuación del plano tangente a la gráfica de $f: \mathbf{R}^2 \to \mathbf{R}$, $(x,y) \mapsto f(x,y)$ en (x_0, y_0), si f fuera suficientemente suave. En $\mathbf{R}^3$, un plano no vertical tiene una ecuación de la forma

$$z = ax + by + c.$$

Si se trata de que sea el plano tangente a la gráfica de f, las pendientes a lo largo de los ejes x y y deben ser iguales a $\partial f/\partial x$ y $\partial f/\partial y$, las tasas de cambio de f respecto a x y y. Así, $a = \partial f/\partial x$, $b = \partial f/\partial y$ (evaluadas en (x_0, y_0)). Finalmente, podemos determinar la constante c a partir del hecho que $z = f(x_0, y_0)$ cuando $x = x_0$, $y = y_0$. Así, obtenemos

$$z = f(x_0, y_0) + \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0), \qquad (1)$$

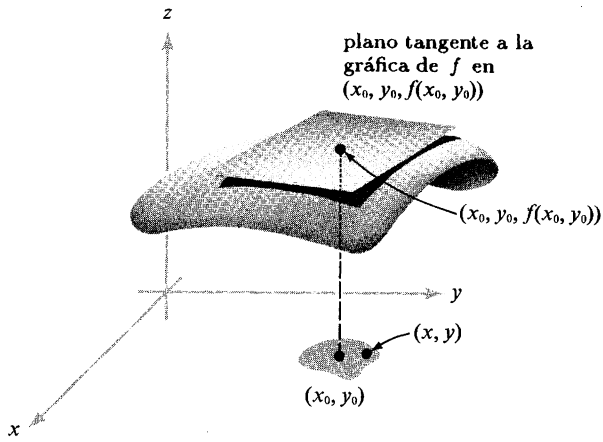


Figura 2.3.4 Para los puntos (x, y) cerca de (x_0, y_0), la gráfica del plano tangente está cerca de la gráfica de f.

que debería ser la ecuación del plano tangente a la gráfica de f en (x_0, y_0), si f es "suficientemente suave" (ver la figura 2.3.4).

Nuestra definición de diferenciabilidad dirá, en efecto, que el plano dado por la ecuación (1) es una "buena" aproximación de f cerca de (x_0, y_0). Para tener una idea de lo que significa una buena aproximación, regresemos por un momento al cálculo en una variable. Si f es diferenciable en el punto x_0, entonces sabemos que

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = f'(x_0).$$

Sea $x = x_0 + \Delta x$ y reescribamos esto como

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0).$$

Usando el límite trivial $\lim_{x \to x_0} f'(x_0) = f'(x_0)$, podemos reescribir la ecuación anterior como

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \to x_0} f'(x_0)$$

o

$$\lim_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right] = 0$$

o

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0.$$

Así, la recta tangente l que pasa por $(x_0, f(x_0))$ con pendiente $f'(x_0)$ está cerca de f en el sentido de que la diferencia entre $f(x)$ y $l(x) = f(x_0) + f'(x_0)(x - x_0)$ se hace cero *aun* al dividirse entre $x - x_0$, cuando x va hacia x_0. Ésta es la idea de "buena aproximación" que adaptaremos a funciones de varias variables, reemplazando la recta tangente por el plano tangente (ver la ecuación (1), anterior).

DEFINICIÓN *Sea* $f: \mathbf{R}^2 \to \mathbf{R}$. *Decimos que* f *es* ***diferenciable en*** (x_0, y_0), *si* $\partial f/\partial x$ *y* $\partial f/\partial y$ *existen en* (x_0, y_0) *y si*

$$\frac{f(x,y) - f(x_0,y_0) - \left[\frac{\partial f}{\partial x}(x_0,y_0)\right](x - x_0) - \left[\frac{\partial f}{\partial y}(x_0,y_0)\right](y - y_0)}{\|(x,y) - (x_0,y_0)\|} \to 0 \qquad (2)$$

cuando $(x, y) \to (x_0, y_0)$. *Esta ecuación expresa el significado que damos cuando decimos que*

$$f(x_0,y_0) + \left[\frac{\partial f}{\partial x}(x_0,y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0,y_0)\right](y - y_0)$$

es una ***buena aproximación*** *a la función* f.

No siempre es fácil usar esta definición para saber si f es diferenciable, pero será fácil usar otro criterio dado en el teorema 9, más adelante.

Hemos usado la idea informal del plano tangente a la gráfica de una función para motivar nuestra definición de diferenciabilidad. Ahora ya estamos preparados para adoptar una definición formal del plano tangente.

DEFINICIÓN *Sea* $f: \mathbf{R}^2 \to \mathbf{R}$ *diferenciable en* $\mathbf{x}_0 = (x_0, y_0)$. *El plano en* $\mathbf{R}^3$ *definido mediante la ecuación (1),*

$$z = f(x_0,y_0) + \left[\frac{\partial f}{\partial x}(x_0,y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0,y_0)\right](y - y_0),$$

se llama ***plano tangente*** *a la gráfica de* f *en el punto* (x_0, y_0).

EJEMPLO 5 *Calcular el plano tangente a la gráfica de* $z = x^2 + y^4 + e^{xy}$ *en el punto* $(1, 0, 2)$.

SOLUCIÓN Aquí usamos la fórmula (1), con $x_0 = 1$, $y_0 = 0$, y $z_0 = f(x_0, y_0) = 2$. Las derivadas parciales son

$$\frac{\partial z}{\partial x} = 2x + ye^{xy} \qquad \text{y} \qquad \frac{\partial z}{\partial y} = 4y^3 + xe^{xy}.$$

En $(1, 0, 2)$, son 2 y 1, respectivamente. Así, por la fórmula (1), el plano tangente es

$$z = 2(x-1) + 1(y-0) + 2, \qquad \text{i.e.} \qquad z = 2x + y. \quad \blacktriangle$$

Escribamos $\mathbf{D}f(x_0, y_0)$ para la matriz renglón

$$\left[\frac{\partial f}{\partial x}(x_0, y_0)\ \frac{\partial f}{\partial y}(x_0, y_0)\right],$$

entonces, la definición de diferenciabilidad afirma que

$$f(x_0, y_0) + \mathbf{D}f(x_0, y_0)\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}$$
$$= f(x_0, y_0) + \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0) \qquad (3)$$

es nuestra buena aproximación a f cerca de (x_0, y_0). (Como antes, "buena" se toma en el sentido de que la expresión (3) difiere de $f(x, y)$ en alguna cantidad pequeña multiplicada por $\sqrt{(x-x_0)^2 + (y-y_0)^2}$.) Decimos que la expresión (3) es la *mejor aproximación lineal* a f cerca de (x_0, y_0).

Ahora estamos preparados para dar una definición de diferenciabilidad para funciones f de $\mathbf{R}^n$ a $\mathbf{R}^m$, usando el análisis anterior como motivación. La derivada $\mathbf{D}f(\mathbf{x}_0)$ de $f = (f_1, \ldots, f_m)$ en un punto $\mathbf{x}_0$ es una matriz con elementos $t_{ij} = \partial f_i / \partial x_j$ evaluada en $\mathbf{x}_0$.*

DEFINICIÓN *Sean U un conjunto abierto en $\mathbf{R}^n$ y $f: U \subset \mathbf{R}^n \to \mathbf{R}^m$ una función dada. Decimos que f es* **diferenciable** *en $\mathbf{x}_0 \in U$ si existen las derivadas parciales de f en $\mathbf{x}_0$ y si*

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{T}(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0, \qquad (4)$$

donde $\mathbf{T} = \mathbf{D}f(\mathbf{x}_0)$ es la matriz cuyos elementos matriciales son $\partial f_i / \partial x_j$ evaluadas en $\mathbf{x}_0$ y $\mathbf{T}(\mathbf{x} - \mathbf{x}_0)$ es el producto de $\mathbf{T}$ con $\mathbf{x} - \mathbf{x}_0$ (considerado como un vector columna). Llamamos a $\mathbf{T}$ **derivada** *de f en $\mathbf{x}_0$.*

Siempre denotaremos la derivada $\mathbf{T}$ de f en $\mathbf{x}_0$ por $\mathbf{D}f(\mathbf{x}_0)$, aunque en algunos libros se denote $df(\mathbf{x}_0)$ y se denomine *diferencial* de f. En el caso $m = 1$, la matriz

*Resulta que es suficiente postular la existencia de *alguna* matriz que dé la mejor aproximación lineal cerca de $\mathbf{x}_0 \in \mathbf{R}^n$, pues de hecho esta matriz es *necesariamente* la matriz cuyo ij-ésimo registro es $\partial f_i / \partial x_j$ (ver la sección 2.7).

$\mathbf{T}$ es precisamente la matriz renglón

$$\left[\frac{\partial f}{\partial x_1}(\mathbf{x}_0) \cdots \frac{\partial f}{\partial x_n}(\mathbf{x}_0)\right].$$

(A veces, cuando hay peligro de confusión, separamos los registros mediante comas.) Al hacer $n = 2$ y colocar el resultado en la ecuación (4), vemos que las condiciones (2) y (4) coinciden. Así, si hacemos $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$, una función f con valores reales, de n variables, es diferenciable en un punto $\mathbf{x}_0$ si

$$\lim_{\mathbf{h}\to 0} \frac{1}{\|\mathbf{h}\|} \left| f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x}_0) h_j \right| = 0,$$

pues

$$\mathbf{Th} = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(\mathbf{x}_0).$$

Para el caso general en que f manda a un subconjunto de $\mathbf{R}^n$ a $\mathbf{R}^m$, la derivada es la matriz de $m \times n$ dada por

$$\mathbf{D}f(\mathbf{x}_0) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_n} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

donde $\partial f_i/\partial x_j$ está evaluada en $\mathbf{x}_0$. A $\mathbf{D}f(\mathbf{x}_0)$ se le llama, de manera correcta, *matriz de las derivadas parciales de f en $\mathbf{x}_0$*.

EJEMPLO 6 *Calcular las matrices de derivadas parciales para* (a) $f(x, y) = (e^{x+y} + y, y^2x)$, (b) $f(x, y) = (x^2 + \cos y, ye^x)$ y (c) $f(x, y, z) = (ze^x, -ye^z)$.

SOLUCIÓN (a) Aquí $f: \mathbf{R}^2 \to \mathbf{R}^2$ está definida por $f_1(x, y) = e^{x+y} + y$ y $f_2(x, y) = y^2x$. Entonces $\mathbf{D}f(x, y)$ es la matriz de 2×2

$$\mathbf{D}f(x, y) = \begin{bmatrix} e^{x+y} & e^{x+y} + 1 \\ y^2 & 2xy \end{bmatrix}.$$

(b) Tenemos

$$\mathbf{D}f(x, y) = \begin{bmatrix} 2x & -\operatorname{sen} y \\ ye^x & e^x \end{bmatrix}.$$

(c) Aquí

$$\mathbf{D}f(x, y, z) = \begin{bmatrix} ze^x & 0 & e^x \\ 0 & -e^z & -ye^z \end{bmatrix}. \quad \blacktriangle$$

DEFINICIÓN *Considerar el caso especial* $f: U \subset \mathbf{R}^n \to \mathbf{R}$. *Aquí* $\mathbf{D}f(\mathbf{x})$ *es una matriz de* $1 \times n$*:*

$$\mathbf{D}f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \end{bmatrix}.$$

Formamos el correspondiente vector $(\partial f/\partial x_1, \ldots, \partial f/\partial x_n)$*, llamado el* ***gradiente*** *de* f *y denotado por* grad f *o* ∇f.

De la definición vemos que para $f: \mathbf{R}^3 \to \mathbf{R}$,

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k},$$

mientras que para $f: \mathbf{R}^2 \to \mathbf{R}$,

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}.$$

El significado geométrico del gradiente se analizará en la sección 2.5. En términos de productos internos, podemos escribir la derivada de f como

$$\mathbf{D}f(\mathbf{x})(\mathbf{h}) = \nabla f(\mathbf{x}) \cdot \mathbf{h}.$$

EJEMPLO 7 Sea $f: \mathbf{R}^3 \to \mathbf{R}$, $f(x, y, z) = xe^y$. Entonces

$$\text{grad } f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = (e^y, xe^y, 0). \quad \blacktriangle$$

EJEMPLO 8 Si $f: \mathbf{R}^2 \to \mathbf{R}$ está dada por $(x, y) \mapsto e^{xy} + \operatorname{sen} xy$, entonces

$$\begin{aligned} \nabla f(x, y) &= (ye^{xy} + y\cos xy)\mathbf{i} + (xe^{xy} + x\cos xy)\mathbf{j} \\ &= (e^{xy} + \cos xy)(y\mathbf{i} + x\mathbf{j}). \end{aligned}$$

En cálculo de una variable se muestra que si f es diferenciable entonces f es continua. En el teorema 8 enunciaremos que esto también se cumple para las funciones diferenciables de varias variables. Como sabemos, hay multitud de funciones que son continuas pero no diferenciables, como $f(x) = |x|$. Antes de enunciar el resultado daremos el ejemplo de una función cuyas derivadas parciales existen en un punto pero que no es continua en ese punto.

EJEMPLO 9 Sea $f: \mathbf{R}^2 \to \mathbf{R}$ definida por

$$f(x,y) = \begin{cases} 1 & \text{si } x = 0 \text{ o si } y = 0 \\ 0 & \text{de no ser así} \end{cases}$$

Como f es constante en los ejes x y y, donde es igual a 1,

$$\frac{\partial f}{\partial x}(0,0) = 0 \qquad \text{y} \qquad \frac{\partial f}{\partial y}(0,0) = 0.$$

Pero f no es continua en $(0,0)$, pues $\lim_{(x,y)\to(0,0)} f(x,y)$ no existe. ▲

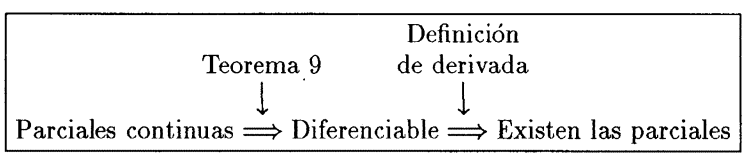
TEOREMA 8 *Sea $f: U \subset \mathbf{R}^n \to \mathbf{R}^m$ diferenciable en $\mathbf{x}_0 \in U$. Entonces f es continua en $\mathbf{x}_0$.*

Consultar la sección 2.7 para ver la demostración.

Como ya vimos, por lo general es fácil decir si existen las derivadas parciales de una función usando nuestro conocimiento de cálculo en una variable. Sin embargo, la definición de diferenciabilidad se ve algo más complicada, y la condición de aproximación requerida en la ecuación (4) parece difícil de verificar. Afortunadamente existe un criterio sencillo, dado en el siguiente teorema, que nos dice cuándo una función es diferenciable.

TEOREMA 9 *Sea $f: U \subset \mathbf{R}^n \to \mathbf{R}^m$. Supongamos que existen todas las derivadas parciales $\partial f_i/\partial x_j$ de f y son continuas en una vecindad de un punto $\mathbf{x} \in U$. Entonces f es diferenciable en $\mathbf{x}$.*

Daremos la demostración en la sección 2.7. Nótese la siguiente jerarquía:



Todos los enunciados recíprocos obtenidos invirtiendo una implicación, son inválidos. (Para un contraejemplo al recíproco de la primera implicación, usar $f(x) = x^2 \operatorname{sen}(1/x)$, $f(0) = 0$; para la segunda, ver el ejemplo 1 de la sección 2.7 o usar el ejemplo 4 de esta sección.)

Una función cuyas parciales existan y sean continuas, se dice que es de *clase* C^1. Así, el teorema 9 dice que *cualquier función* C^1 *es diferenciable.*

EJEMPLO 10 *Sea*

$$f(x, y) = \frac{\cos x + e^{xy}}{x^2 + y^2}.$$

Mostrar que f *es diferenciable en todos los puntos* $(x, y) \neq (0, 0)$.

SOLUCIÓN Observar que las derivadas parciales

$$\frac{\partial f}{\partial x} = \frac{(x^2 + y^2)(ye^{xy} - \operatorname{sen} x) - 2x(\cos x + e^{xy})}{(x^2 + y^2)^2}$$

$$\frac{\partial f}{\partial y} = \frac{(x^2 + y^2)xe^{xy} - 2y(\cos x + e^{xy})}{(x^2 + y^2)^2}$$

son continuas excepto cuando $x = 0$ y $y = 0$ (por los resultados en la sección 2.2). ▲

EJERCICIOS

1. Hallar $\partial f/\partial x$, $\partial f/\partial y$ si

(a) $f(x, y) = xy$

(b) $f(x, y) = e^{xy}$

(c) $f(x, y) = x \cos x \cos y$

(d) $f(x, y) = (x^2 + y^2) \log(x^2 + y^2)$

2. Evaluar las derivadas parciales $\partial z/\partial x$, $\partial z/\partial y$ para las funciones dadas en los puntos indicados.

(a) $z = \sqrt{a^2 - x^2 - y^2}$; $(0, 0)$, $(a/2, a/2)$

(b) $z = \log \sqrt{1 + xy}$; $(1, 2)$, $(0, 0)$

(c) $z = e^{ax} \cos(bx + y)$; $(2\pi/b, 0)$

3. En cada caso siguiente, hallar las derivadas parciales $\partial w/\partial x$, $\partial w/\partial y$.

(a) $w = xe^{x^2+y^2}$

(b) $w = \dfrac{x^2 + y^2}{x^2 - y^2}$

(c) $w = e^{xy} \log(x^2 + y^2)$

(d) $w = x/y$

(e) $w = \cos ye^{xy} \operatorname{sen} x$

4. Mostrar que cada una de las funciones siguientes es diferenciable en cada punto de su dominio. Decir cuáles de las funciones son C^1.

(a) $f(x,y) = \dfrac{2xy}{(x^2+y^2)^2}$ (b) $f(x,y) = \dfrac{x}{y} + \dfrac{y}{x}$

(c) $f(r,\theta) = \frac{1}{2} r \operatorname{sen} 2\theta, r > 0$ (d) $f(x,y) = \dfrac{xy}{\sqrt{x^2+y^2}}$

(e) $f(x,y) = \dfrac{x^2 y}{x^4+y^2}$

5. Hallar la ecuación del plano tangente a la superficie $z = x^2 + y^3$ en $(3,1,10)$.

6. Usando las funciones respectivas en el ejercicio 1, calcular el plano tangente a la gráfica en los puntos indicados.

(a) $(0,0)$ (b) $(0,1)$ (c) $(0,\pi)$ (d) $(0,1)$

7. Calcular la matriz de derivadas parciales de las funciones siguientes:

(a) $f: \mathbf{R}^2 \to \mathbf{R}^2, f(x,y) = (x,y)$

(b) $f: \mathbf{R}^2 \to \mathbf{R}^3, f(x,y) = (xe^y + \cos y, x, x + e^y)$

(c) $f: \mathbf{R}^3 \to \mathbf{R}^2, f(x,y,z) = (x + e^z + y, yx^2)$

(d) $f: \mathbf{R}^2 \to \mathbf{R}^3, f(x,y) = (xye^{xy}, x \operatorname{sen} y, 5xy^2)$

8. Calcular la matriz de derivadas parciales de

(a) $f(x,y) = (e^x, \operatorname{sen} xy)$ (b) $f(x,y,z) = (x-y, y+z)$

(c) $f(x,y) = (x+y, x-y, xy)$ (d) $f(x,y,z) = (x+z, y-5z, x-y)$

9. ¿Dónde cruza al eje z el plano tangente a $z = e^{x-y}$ en $(1,1,1)$?

10. ¿Por qué podrían llamarse las gráficas de $f(x,y) = x^2+y^2$ y $g(x,y) = -x^2-y^2+xy^3$ "tangentes" en $(0,0)$?

11. Sea $f(x,y) = e^{xy}$. Mostrar que $x(\partial f/\partial x) = y(\partial f/\partial y)$.

12. Usar la expresión (1) en esta sección para aproximar una función adecuada $f(x,y)$ y a partir de ahí estimar lo siguiente:

(a) $(0.99e^{0.02})^8$

(b) $(0.99)^3 + (2.01)^3 - 6(0.99)(2.01)$

(c) $\sqrt{(4.01)^2 + (3.98)^2 + (2.02)^2}$

13. Calcular los gradientes de las funciones siguientes:

(a) $f(x,y,z) = x \exp(-x^2 - y^2 - z^2)$ (Notar que $\exp u = e^u$.)

(b) $f(x,y,z) = \dfrac{xyz}{x^2+y^2+z^2}$ (c) $f(x,y,z) = z^2 e^x \cos y$

14. Calcular el plano tangente a $(1,0,1)$ para cada una de las funciones en el ejercicio 13. (La solución sólo a la parte (c) está en la Guía de estudio de este libro.)

15. Hallar la ecuación del plano tangente a $z = x^2 + 2y^3$ en $(1,1,3)$.

16. Calcular $\nabla h(1,1,1)$ si $h(x,y,z) = (x+z)e^{x-y}$.

17. Sea $f(x,y,z) = x^2 + y^2 - z^2$. Calcular $\nabla f(0,0,1)$.

18. Evaluar el gradiente de $f(x,y,z) = \log(x^2 + y^2 + z^2)$ en $(1,0,1)$.

***19.** Describir todas las funciones Hölder-continuas con $\alpha > 1$ (ver el ejercicio 21, sección 2.2). [IDEA: ¿Cuál es la derivada de dicha función?]

***20.** Suponer que $f: \mathbf{R}^n \to \mathbf{R}^m$ es una transformación lineal. ¿Cuál es la derivada de f?

2.4 PROPIEDADES DE LA DERIVADA

En cálculo elemental aprendimos cómo diferenciar sumas, productos, cocientes y funciones compuestas. Ahora generalizamos estas ideas a funciones de varias variables, prestando particular atención a la diferenciación de funciones compuestas. La regla para diferenciación de composiciones, llamada regla de la cadena, adquiere una forma más profunda en el caso de funciones de varias variables que en las de una variable. Así, por ejemplo, si f es una función con valores reales en una variable, escrita como $z = f(y)$, y y es una función de x, que escribimos $y = g(x)$, entonces z resulta una función de x mediante la sustitución, a saber, $z = f(g(x))$, y tenemos la conocida fórmula

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = f'(g(x))g'(x).$$

Si f es una función con valores reales en tres variables u, v y w, escrita en la forma $z = f(u,v,w)$, y las variables u, v y w son cada una funciones de x, $u = g(x)$, $v = h(x)$, y $w = k(x)$, entonces al sustituir $g(x)$, $h(x)$ y $k(x)$ por u, v y w, expresamos z como función de x: $z = f(g(x), h(x), k(x))$. En este caso la regla de la cadena es:

$$\frac{dz}{dx} = \frac{\partial z}{\partial u}\frac{du}{dx} + \frac{\partial z}{\partial v}\frac{dv}{dx} + \frac{\partial z}{\partial w}\frac{dw}{dx}.$$

Uno de los objetivos de esta sección es explicar dichas fórmulas en detalle.

Comenzamos con las reglas de diferenciación para sumas, productos y cocientes.

TEOREMA 10

(i) ***Regla del múltiplo constante.*** *Sea* $f: U \subset \mathbf{R}^n \to \mathbf{R}^m$ *diferenciable en* $\mathbf{x}_0$ *y sea* c *un número real. Entonces* $h(\mathbf{x}) = cf(\mathbf{x})$ *es diferenciable en* $\mathbf{x}_0$ *y*

$$\mathbf{D}h(\mathbf{x}_0) = c\mathbf{D}f(\mathbf{x}_0) \qquad \textit{(igualdad de matrices).}$$

(ii) ***Regla de la suma.*** *Sean* $f: U \subset \mathbf{R}^n \to \mathbf{R}^m$ *y* $g: U \subset \mathbf{R}^n \to \mathbf{R}^m$ *diferenciables en* $\mathbf{x}_0$. *Entonces* $h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ *es diferenciable en* $\mathbf{x}_0$ *y*

$$\mathbf{D}h(\mathbf{x}_0) = \mathbf{D}f(\mathbf{x}_0) + \mathbf{D}g(\mathbf{x}_0) \qquad \textit{(suma de matrices).}$$

(iii) ***Regla del producto.*** *Sean* $f: U \subset \mathbf{R}^n \to \mathbf{R}$ *y* $g: U \subset \mathbf{R}^n \to \mathbf{R}$ *diferenciables en* $\mathbf{x}_0$ *y sea* $h(\mathbf{x}) = g(\mathbf{x})f(\mathbf{x})$. *Entonces* $h: U \subset \mathbf{R}^n \to \mathbf{R}$ *es diferenciable en* $\mathbf{x}_0$ *y*

$$\mathbf{D}h(\mathbf{x}_0) = g(\mathbf{x}_0)\mathbf{D}f(\mathbf{x}_0) + f(\mathbf{x}_0)\mathbf{D}g(\mathbf{x}_0).$$

(Notar que cada lado de esta ecuación es una matriz de $1 \times n$*; un producto más general se presenta en el ejercicio 25 al final de esta sección.)*

(iv) ***Regla del cociente.*** *Con las mismas hipótesis que en la regla* (iii), *sea* $h(\mathbf{x}) = f(\mathbf{x})/g(\mathbf{x})$ *y suponer que* g *nunca es cero en* U. *Entonces* h *es diferenciable en* $\mathbf{x}_0$ *y*

$$\mathbf{D}h(\mathbf{x}_0) = \frac{g(\mathbf{x}_0)\mathbf{D}f(\mathbf{x}_0) - f(\mathbf{x}_0)\mathbf{D}g(\mathbf{x}_0)}{[g(\mathbf{x}_0)]^2}.$$

DEMOSTRACIÓN Las demostraciones de las reglas (i) a (iv) se desarrollan casi de la misma manera que en el caso de una variable, con sólo una ligera diferencia en la notación. Probaremos las reglas (i) y (ii), dejando las demostraciones de las reglas (iii) y (iv) para el ejercicio 29.

(i) Para mostrar que $\mathbf{D}h(\mathbf{x}_0) = c\mathbf{D}f(\mathbf{x}_0)$, debemos mostrar que

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\|h(\mathbf{x}_0) - h(\mathbf{x}_0) - c\mathbf{D}f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0,$$

esto es, que

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \frac{\|cf(\mathbf{x}) - cf(\mathbf{x}_0) - c\mathbf{D}f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0,$$

(ver la ecuación (4) de la sección 2.3). Esto es cierto pues f es diferenciable y la constante c puede factorizarse (ver el teorema 3(i), sección 2.2).

(ii) Por la desigualdad del triángulo, podemos escribir

$$\begin{aligned}
&\frac{\|h(\mathbf{x}) - h(\mathbf{x}_0) - [\mathbf{D}f(\mathbf{x}_0) + \mathbf{D}g(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} \\
&\quad = \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - [\mathbf{D}f(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0) + g(\mathbf{x}) - g(\mathbf{x}_0) - [\mathbf{D}g(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} \\
&\quad \leq \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - [\mathbf{D}f(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} \\
&\quad + \frac{\|g(\mathbf{x}) - g(\mathbf{x}_0) - [\mathbf{D}g(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|},
\end{aligned}$$

y cada término tiende a 0 conforme $\mathbf{x} \to \mathbf{x}_0$. Por lo tanto, se cumple la regla (ii). ■

EJEMPLO 1 *Verificar la fórmula para* $\mathbf{D}h$ *en la regla* (iv) *del teorema 10 con* $f(x, y, z) = x^2 + y^2 + z^2$ *y* $g(x, y, z) = x^2 + 1$.

SOLUCIÓN Aquí

$$h(x, y, z) = \frac{x^2 + y^2 + z^2}{x^2 + 1},$$

de modo que por diferenciación directa

$$\begin{aligned}
\mathbf{D}h(x, y, z) &= \left[\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}, \frac{\partial h}{\partial z}\right] \\
&= \left[\frac{(x^2 + 1)2x - (x^2 + y^2 + z^2)2x}{(x^2 + 1)^2}, \frac{2y}{x^2 + 1}, \frac{2z}{x^2 + 1}\right] \\
&= \left[\frac{2x(1 - y^2 - z^2)}{(x^2 + 1)^2}, \frac{2y}{x^2 + 1}, \frac{2z}{x^2 + 1}\right].
\end{aligned}$$

Por la regla (iv), obtenemos

$$\mathbf{D}h = \frac{g\mathbf{D}f - f\mathbf{D}g}{g^2} = \frac{(x^2 + 1)[2x, 2y, 2z] - (x^2 + y^2 + z^2)[2x, 0, 0]}{(x^2 + 1)^2}.$$

que es lo mismo que obtuvimos directamente. ▲

Como ya mencionamos anteriormente, es en la diferenciación de funciones compuestas que encontraremos aparentes alteraciones substanciales de la fórmula del cálculo de una variable. Sin embargo, si usamos la notación $\mathbf{D}$, esto es, notación matricial, la regla de la cadena para funciones de varias variables se parece a la regla para una variable.